6-28-2020

Data-Driven and Model-Based Methods with Physics-Guided Machine Learning for Damage Identification

Zhiming Zhang
Louisiana State University and Agricultural and Mechanical College

DATA-DRIVEN AND MODEL-BASED METHODS WITH PHYSICS-GUIDED MACHINE LEARNING FOR DAMAGE IDENTIFICATION

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Civil & Environmental Engineering

by
Zhiming Zhang
B.S., Liaoning Technical University, 2010
M.S., Southwest JiaotongUniversity, 2013
August 2020

Acknowledgments

This dissertation would not be possible without several contributions. It is a pleasure to thank my advisor, Dr. Chao Sun. It would be difficult to imagine the completion of this dissertation without his considerable and consistent support, helpful suggestions, and valuable encouragement on my research during my Ph.D. study. It was a wonderful experience working with such a knowledgeable and inspiring scholar.

I would like to appreciate sincerely the support from the rest of my committee members: Dr. Steve C.S. Cai, Dr. Ayman Okeil, Dr. Beibei Guo, and Dr. Yongcheol Lee. They offered beneficial suggestions on my research, which allows me to further improve the quality of my dissertation.

It is a pleasure also to thank my fellow members of Dr. Sun's group for providing me with a pleasant working environment and helping me whenever necessary. It was a wonderful experience interacting with them in the past four years.

The financial support for this research from the Louisiana Transportation Research Center (LTRC) and Louisiana Board of Regents is fully acknowledged.

This dissertation is dedicated to my wife and parents for their love, support, and encouragement.

Table of Contents

Abstract

Structural health monitoring (SHM) has been widely used for structural damage diagnosis and prognosis of a wide range of civil, mechanical, and aerospace structures. SHM methods are generally divided into two categories: (1) model-based methods; (2) data-driven methods. Compared with data-driven SHM, model-based methods provide an updated physics-based numerical model that can be used for damage prognosis when long-term data is available. However, the performance of model-based methods is susceptible to modeling error in establishing the numerical model, which is usually unavoidable due to model simplification and omission. The major challenge of data-driven SHM methods lies in data insufficiency, e.g., lack of data covering as many as possible damage states, especially for large-scale structures. Hence, multi-site damage identification using data-driven methods can be more challenging as pattern recognition theoretically requires sufficient data from each damage scenario.

The main objectives of this dissertation are to: (1) integrate model-based and data-driven SHM methods so that their shortcomings can be weakened while their respective merits can be preserved when implementing damage identification; (2) improve the accuracy of data-driven methods for multi-site damage identification with limited measured data.

To achieve the first research objective, physics-guided machine learning (PGML) is proposed to improve the performance of pattern recognition in data-driven SHM with insufficient measured data. The results of model-based SHM (i.e., FE model updating) are taken as an implicit representation of physics underlying the monitored structure, which is

incorporated into the learning process of a neural network model with the physics guidance introduced into the loss function. In addition to PGML, transfer learning (TL) is used to bridge the gap between the numerical and experimental domains of SHM. The distribution difference and manifold discrepancy between the two domains is minimized through TL as a means of domain adaptation.

To improve the performance of multi-site damage identification in data-driven SHM, multi-label classification (MLC) and constrained independent component analysis(cICA) methods are applied to investigate the correlations between damage cases sharing common damaged sites. Finally, as a case study, a two-step strategy of identifying structural damage of offshore wind turbines via FE model updating is proposed.

Chapter 1. Introduction

1.1. Introduction

Structural health monitoring (SHM) refers to structural identification and damage evaluation of engineering structures such as buildings and bridges [1]. In SHM, damage means changes to the structural properties including material and/or geometric properties, boundary conditions, and structural connectivity, which can reduce the structural performance considerably. The SHM process includes structural observation over a long period of time using measured responses with an array of sensors (e.g., accelerometers), extracting damage sensitive metrics and features (e.g, natural frequencies and mode shapes), evaluating the current state of structural health condition, and finally estimating the remaining life of monitored structure. In addition to long term structural monitoring and evaluation with a focus on structural aging and degradation under service and operational loads, SHM is also used during extreme events, such as earthquakes and hurricanes, for rapid realtime evaluation of structural integrity and functionality.

According to the algorithm used for structural identification and damage evaluation given measured data, SHM approaches can be divided into two categories: model-based and data-driven approaches. Model-based SHM is mainly implemented through Finite element (FE) model updating, which is a procedure calibrating the FE model parameters of the target structure such that its outputs match the experimental measurements [2]. FE model updating is solved by tuning the model parameters in a constrained optimization problem with an objective function defined as the discrepancy between the model outputs and experimental measurements. The parameters in FE model updating to be optimized

are usually selected in accordance with the SHM requirements, which can be parameters of material/geometric properties, support stiffness, etc. With structural parameters tuned in FE model updating, structural damage can be evaluated via comparing the current structural parameters with that of the intact condition. Moreover, the updated FE model can be used for damage prognosis when long-term monitoring is feasible. While model-based methods are extensively studied for damage diagnosis, they are susceptible to modeling uncertainty or modeling error due to model idealizations and simplifications that make the FE model updating results deviate from the truth [3]. Modeling uncertainty or modeling error has been widely recognized as one major challenge in FE model updating for structural identification and damage detection. It renders model updating inherently ineffective in converging to the real structural model because of the physical bias present in establishing the numerical model.

Differing from model based approaches, data-driven SHM detects damage occurrence and identifies its location and severity through pattern recognition and machine learning methods using damage-sensitive features extracted from collected structural responses [4]. This paradigm contains four components: (1) operational evaluation; (2) data acquisition, cleansing, and normalization; (3) feature extraction and selection; (4) pattern recognition for structural damage evaluation. Data-driven SHM faces the challenge of data insufficiency that renders the learned model limited in identifying damage scenarios that are not contained in the training data. On the other hand, while existing studies in this field presented encouraging results for single-site damage identification, limited research effort has been devoted to identifying multi-site damage due to its complexity. Efficient features for single-site damage identification may lose their effectiveness when multi-site

damage occurs [5]. Compared with model-based methods, data-driven SHM does not require a numerical model and thus can avoid the modeling uncertainty/error issues. It also has advantages in dealing with other uncertainties through data cleansing, compression, and pattern recognition.

1.2. Motivation and Objectives of This Study

Section 1.1 introduces the merits of model-based and data-driven SHM methods as well as their shortcomings. It shows that the model-based FE model updating and data-driven pattern recognition are complementary SHM methods. Hence, it becomes attractive to combine these two methods synergistically such that they can work in a collaborative manner. This combination has the promise to overcome the shortcomings of current SHM methods and enhance their performance. To this end, this dissertation covers the following topics in sufficient details: (1) bridging the gap between the model-based and data-driven SHM methods; (2) improving the performance of model-based SHM via incorporating pattern recognition; (3) improving the data-driven SHM performance via physics guidance; (4) bridging the gap between numerical models and physical structures in SHM using transfer learning.

In addition, data-driven multi-site damage identification is investigated with the target of improving the accuracy of damage evaluation with limited data quantity and quality and relieving the data insufficiency issue in data-driven SHM. The research objectives of this dissertation are listed as follows:

(1) Resolving the data insufficiency problem in data-driven SHM. Data insufficiency can cause overfitting in pattern recognition and thus adversely affects the damage identi-

fication performance. On one hand, a physics-based model can provide knowledge in regions not sufficiently covered by measured data. Hence the physics-related consistency of the SHM outcome can be improved if the physics guidance is incorporated into the learning process of pattern recognition. On the other hand, a numerical model of the target structure can provide sufficient data of almost all damage conditions. Hence a machine learning model with sufficient generality can be learned in the numerical domain. Then the knowledge learned in the numerical domain can be transfered to the experimental domain for damage evaluation through domain adaptation.

(2) Resolving the modeling uncertainty/error issue in model-based SHM. Modeling uncertainty is unavoidable when establishing the numerical model and leads to biased results of model updating and the following damage evaluation. Pattern recognition in data-driven SHM does not necessitate a numerical model and thus can obviate modeling uncertainty. The pattern recognition results can help regularize the process of FE model updating (though with an imperfect model) and thus reduce the influence of considerable modeling uncertainty.

(3) Improving the performance of multi-site damage detection in data-driven SHM. Multi-site damage identification faces the challenge of data insufficiency in pattern recognition. Traditional pattern recognition methods such as multiclass classification require sufficient data for all possible damage scenarios of interest, which is unachievable for most engineering structures. However, if the underlying physical intercorrelation between damage cases sharing damaged sites can be considered, the accuracy of multi-site damage identification can be significantly improved. This improvement will in turn reduce the burden of data sufficiency in multi-site damage identifica-

tion.

1.3. Structure of the Dissertation

This thesis consists of seven chapters. Chapter 1 introduces the SHM concept and methods, research motivation and objectives, and the structure of this dissertation.

Chapter 2 proposes a physics-guided machine learning (PGML) method to improve the performance of data-driven SHM as well as model-based SHM. PGML uses observed feature data with correct labels as well as the physical model output of unlabeled instances. In this study, PGML is realized with a physics-guided neural network (PGNN). The original modal-property based features are extended with the damage identification result of FE model updating. A physics-based loss function is designed and introduced to evaluate the discrepancy between the neural network (NN) model output and that of FE model updating. With the guidance from the scientific knowledge contained in FE model updating, the learned NN model has the potential to improve the generality and scientific consistency of the damage detection results. The proposed methodology is validated by a numerical case study on a steel pedestrian bridge model and an experimental study on a 3-story building model.

Chapter 3 proposes a methodology that applies pattern recognition methods to guide Bayesian model updating (BMU) and supervise the identification of structural damage. In detail, the domain adaptation technique realized by transfer learning (TL) is used to bridge the gap between the biased numerical model and the real structure and to guide the model updating process. Numerical and experimental studies have been implemented

to evaluate the performance of domain adaptation in identifying the correct damage locations. It is found that TL-guided BMU is advantageous over the traditional method in identifying damage severities when modeling error exists.

Chapter 4 proposes a new method to locate multi-site structural damage using a data-driven multi-label classification (MLC) method. Differing from the multiclass classification (MCC) scheme that sets each damage case as a category, the MLC method denotes each damage case with multiple labels, with each label indicating the occurrence of damage at a certain location. The MLC method considers the physical correlation between damage cases sharing damage locations which is neglected in the MCC method. This study uses the instance differentiation algorithm to implement MLC. Damage identification results obtained based on numerical and experimental data indicate that the MLC can identify multi-site damage with good accuracy, even for damage cases that are not covered by the training dataset. Through comparison, it is demonstrated that the MLC method outperforms the MCC and binary classification methods for multi-site damage identification. Moreover, MLC preserves generality when tested on unobserved multi-site damage cases.

Chapter 5 presents another solution to the multi-site structural damage identification problem using a data-driven method and constrained independent component analysis (cICA). This method extracts damage-sensitive features from the ICA outcome of the structural responses under certain excitations. The information of structural damage contained in the response is compacted into the mixing matrix by enforcing identical independent components to that of the intact structure. Hence, the cICA can significantly reduce the feature dimension and preserve all the valuable information of damage. A case study

indicates that the mixing matrix elements, when used as damage features, can distinguish multi-site damage cases from single-site damage cases and accurately locate the single-site damage. Furthermore, the mixing matrix columns of multi-site damage cases exhibit distinct correlation with that of the corresponding single-site damage cases. As a result, the proposed method can progressively and accurately locate the multi-site structural damage. Moreover, the proposed method has the potential to identify multi-site damage without requiring the corresponding multi-site damage data as a reference. This relieves the burden of data incompleteness when using data-driven damage identification methods and pattern recognition.

Chapter 6 proposes a two-step strategy of identifying structural damage of off-shore wind turbines via Finite Element (FE) model updating. The first step of the proposed method focuses on global monitoring and identifies the damaged component(s) using a verified global mathematical model of the offshore wind turbine under operational wind and wave loading. Damaged components include the tower, certain single or multiple blade(s), and/or the foundation. Natural frequencies and mode shapes obtained from operational modal analysis are used to formulate the objective function for deterministic model updating. With the damaged component(s) identified in the first step, the second step of the method focuses on detecting the exact damage location on the damaged component(s) and quantifying the corresponding severities. To this end, an FE model of a baseline 5 megawatts monopile offshore wind turbine is established using ANSYS [6]. Wind and wave loading are applied to simulate the structural responses under different damage scenarios. In additional to natural frequencies and mode shapes, mode shape curvatures obtained in the second step are included in the objective function. Results indicate that the proposed

methodology can correctly identify most structural damage of the offshore wind turbines using operational data.

Finally, Chapter 7 presents the critical findings of this dissertation and proposes ideas and thoughts for future research on SHM topics related to this dissertation.

Chapter 2. Structural Damage Identification via Physics-Guided Machine Learning: A Methodology Integrating Pattern Recognition with Finite Element Model Updating

2.1. Introduction

Structural health monitoring (SHM) approaches mainly fall into two categories: data-driven and model-based approaches. Data-driven approaches detect damage occurrence and identifies its location and severity through pattern recognition and machine learning methods using damage-sensitive features extracted from collected structural responses [7]. Compared with model-based methods of SHM, a data-driven approach avoids building and validating a numerical model [8, 9] and has the potential of identifying structural damage despite the operational and environmental influence such as traffic loading across bridges, temperature variations, and wind and moisture effects [10, 11]. Moreover, it automatically accommodates the uncertainty that originates from measuring variability [9]. While data-driven methods have aforementioned advantages over model-based methods, a big challenge of data-driven SHM is the availability of sufficient training data with correct labels for learning a statistical model with satisfactory accuracy and generality. Specifically, damage localization in data-driven SHM is a supervised learning problem setting the potential damage locations as the target class labels of a machine learning classifier [5]. This learning process requires training data from both undamaged and damaged conditions. However, such data especially that of the damaged cases will always be lacked for large and valuable structures, e.g., long span bridges and offshore wind turbines,

Reprinted from Zhiming Zhang and Chao Sun, Structural Damage Identification via Physics-Guided Machine Learning: A Methodology Integrating Pattern Recognition with Finite Element Model Updating, Structural Health Monitoring. Copyright ©[2020] (Sage). DOI: [10.1177/1475921720927488].

which enlarges the over-fitting probability of the learned diagnostic model. The lack of data is probably the greatest challenge in applying pattern recognition and machine learning methods in SHM [9, 12].

Differing from data-driven SHM approaches, model-based approaches evaluate structural condition through updating a representative physics-based model of the target structure, such as a finite element (FE) model, by minimizing the discrepancy of its predictions from the measured data [12, 5]. Compared with data-driven approaches, a model-based approach provides a calibrated physics-based numerical model that can be used for damage prognosis. However, a critical barrier limiting in-practice application of FE model updating is the modeling error that originates from model simplification and omission. In model-based SHM, modeling error renders the updated model biased from the real structure, which leads to challenge in structural parameter estimation [13], structural damage detection [14] and predicting structural features and responses [15].

Considering that both data-driven and model-based approaches have critical shortcomings and that their merits are complementary, it would be attractive if they can be synergistically integrated in SHM so that their merits get preserved and their shortcomings become less critical. To this end, the present study proposes integrating pattern recognition with FE model updating in SHM via physics-guided machine learning (PGML). PGML leverages measured data with correct labels as well as the scientific knowledge contained in the physics-based model (FE model in SHM), so that the model predictions are consistent with the scientific principles behind the physics-based model and maintain sufficient accuracy on the labeled data. It has the potential of improving the generality of the learned model and scientific consistency of its predictions even when

representative labeled samples are very limited [16]. PGML has been broadly applied in many areas such as climate pattern discovery [17], turbulence modeling [18], material science [19], quantum chemistry [20], etc.

Based on the literature review, a new method using PGML for structural damage identification is proposed and evaluated in the present study. PGML is realized with a multi-layer perceptron (MLP) neural network (NN) model, that is physics-guided neural networks (PGNN) [21], through extending the original modal-property based feature with the damage localization output of FE model updating and incorporating a physics-based loss function. The physics-based loss function evaluates the discrepancy between the output of the NN model and that of FE model updating. With physics guidance from the updated FE model, the learned NN model generalizes well to the unseen test data. Moreover, it is shown that errors in damage locations and severities can be significantly reduced by integrating the results of damage localization with PGNN into FE model updating. The efficiency of the proposed methodology in structural damage localization is validated numerically using a steel pedestrian bridge model [22] and experimentally using measured data from a 3-story building model [23].

2.2. Methodology

To incorporate physics into data-driven SHM and realize PGML for structural damage evaluation, the present study uses the FE model as an implicit representation of scientific knowledge underlying the monitored structure and incorporates the output of FE model updating into the NN model setup and learning. This section establishes the framework of PGML for structural damage localization, which contains two major steps: (1)

extending the original feature vector with the output of FE model updating; (2) designing a physics-based loss function that integrates the scientific knowledge underlying the FE model into the NN model learning process. This section first introduces the method of FE model updating used in this study, then describes the two major steps of PGML and introduces how to implement PGML in PyTorch, a framework of deep learning.

2.2.1. FE Model Updating

For a monitored structure, the stiffness matrix $\mathbf{K}$ can be formulated as:

$$\mathbf{K} = \mathbf{K}_0 + \sum_{i=1}^{n_\alpha} \alpha_i \mathbf{K}_i \tag{2.1}$$

in which $\mathbf{K}_0$ is the sum of known substructural stiffness matrices prior to model updating; $\mathbf{K}_i$ is the nominal stiffness matrix of substructure i with unknown stiffness; α_i is the coefficient corresponding to $\mathbf{K}_i$; n_α is the number of substructures with unknown stiffness. Hence, $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_{n_\alpha}]$ containing all unknown stiffness coefficients is the target of FE model updating.

Measured modal properties including natural frequencies and mode shapes are usually used to formulate the objective function of model updating. This study adopts the formulation with the eigen-frequency and mode shape differences [24]. That is:

$$\begin{aligned} &L(\boldsymbol{\alpha}) = \\ &\sum_{i=1}^{n_\mathrm{m}} \left\{ \left(\frac{\lambda_i^\mathrm{e} - \lambda_i(\boldsymbol{\alpha})}{\lambda_i^\mathrm{e}} \cdot w_{\lambda_i} \right)^2 + ||\mathbf{Q}_i \left(\boldsymbol{\Phi}_i^\mathrm{e} - \boldsymbol{\Phi}_i^\mathrm{m}(\boldsymbol{\alpha}) \right) \cdot w_{\Phi_i}||_2^2 \right\} \end{aligned} \tag{2.2}$$

in which n_m is the number of measured modes in dynamic tests; λ_i^e is the experimentally measured eigen-frequency of the i^th mode; $\lambda_i(\boldsymbol{\alpha})$ is the evaluated value of λ_i from

the FE model using a certain value of $\boldsymbol{\alpha}$; $\boldsymbol{\Phi}_i^{\mathrm{e}}$ is the measured mode shape of the i^{th} mode; $\boldsymbol{\Phi}_i^{\mathrm{m}}(\boldsymbol{\alpha})$ is the evaluated $\boldsymbol{\Phi}_i$ at the measured DOFs using $\boldsymbol{\alpha}$. $\mathbf{Q}_i$ is the selection matrix; w_{λ_i} and w_{Φ_i} are the weighting factors of the eigen-frequency and mode shape respectively.

The Levenberg-Marquardt algorithm is selected for optimization and is implemented using the "lsqnonlin" solver in MATLAB. The Jacobian derivative is used to determine the local search direction at each iteration. A number of runs, for example 50, are implemented with random starting points, and the solution yielding the least objective function is selected as the final solution.

2.2.2. Feature Extension

Modal properties, including the natural frequencies and mode shapes, are widely used in data-driven SHM in deriving damage-sensitive features and designing objective function in FE model updating. This study uses the normalized frequency change ratio (NFCR) and the change of mode shapes $\mathrm{d}\Phi$ of the first several modes as features (X) for damage detection [25]. That is:

$$X = \{\mathrm{NFCR}; \mathrm{d}\Phi\} \tag{2.3}$$

$\mathrm{d}\Phi$ is the difference of mode shapes between damaged and intact cases. NFCR of a certain mode is calculated as the normalized fractional frequency change (FFC), that is:

$$\mathrm{NFCR}_i = \frac{\mathrm{FFC}_i}{\sum_{j=1}^{N} \mathrm{FFC}_j} \tag{2.4}$$

in which N is the number of modes selected for SHM purpose. FFC for the i^{th} mode is expressed as:

$$\mathrm{FFC}_i = \frac{f_{\mathrm{u}i} - f_{\mathrm{d}i}}{f_{\mathrm{u}i}} \tag{2.5}$$

in which f_{ui} and f_{di} are the i^{th} mode frequencies of the structure in undamaged and damaged states, respectively.

After preparing data with features X and labels y, a standard NN model would be $f_{\text{NN}} : X \rightarrow y$ that yields estimated labels $\hat{y}$. Alternatively, FE model updating using the measured modal properties will yield damage severities (i.e., z_{mu}) at interested sites and thus recommend the most probable damage locations. In this study, y_{mu} is set as the location with the most severe damage in z_{mu}. It is noted that y_{mu} may not be a correct representation of the structural damage distribution due to simplification and/or idealization in FE model establishment and updating. To address this issue, PGML integrates the output of a physics-based model into the original feature data, so that both information from measured data and physics can be leveraged in learning a model [26]. For example, in an NN model, the hidden layers can extract complex features from the extended feature input so that the insufficiency of the physics-based model can be complemented. Then we have the extended feature for damage detection in SHM as follows:

$$X_{\text{ext}} = \{X; y_{\text{mu}}\} \tag{2.6}$$

X_{ext} will be used as input of the PGNN model in this study.

2.2.3. Physics-based Loss Function

A standard NN model using the extended feature X_{ext} as input aims to minimize the data loss calculated on the labeled training data as well as the model complexity that is expressed as the regularization terms of the model parameter norms. So the loss function is defined as:

$$Loss = L_{\text{d}}(z_{\text{f}}, y) + \lambda_{\text{R}} R(f) \tag{2.7}$$

in which L_{d} denotes the data loss; z_{f} is the output scores of the NN model f; y denotes the target labels of the labeled data; $R(f)$ measures the model's complexity or structural loss; λ_{R} is the regularization parameter. For a certain input data x_i the output of an NN model $z_{\mathrm{f}}(x_i)$ contains the scores of each possible class, that is $z_{\mathrm{f}}(x_i) = [z_{\mathrm{f}}^1(x_i), z_{\mathrm{f}}^2(x_i), ..., z_{\mathrm{f}}^c(x_i)]$ in which c is the number of classes or number of locations of interest in SHM. For a multiclass problem in data-driven structural damage localization, the cross-entropy loss is used to evaluate the classification performance [27]. Then the data loss of x_i is:

$$L_{\mathrm{d}}\left(z_{\mathrm{f}}(x_i), y_i\right) = -z_{\mathrm{f}}^{y_i}(x_i) + \log\left(\sum_{j=1}^{c} \exp\left(z_{\mathrm{f}}^{j}(x_i)\right)\right) \tag{2.8}$$

The ℓ_1 and ℓ_2 norms of the network weights W are regularized to control the model complexity. That is:

$$\lambda_{\mathrm{R}} R(W) = \lambda_1 \|W\|_1 + \lambda_2 \|W\|_2 \tag{2.9}$$

in which λ_1 and λ_2 are the regularization parameters.

Considering the limitations of standard machine learning procedures for the SHM problem with limited labeled data, physics-based loss functions or scientific inconsistency L_{p} are introduced into the PGML to guide the learning process for physically consistent solutions. Then the loss function becomes:

$$Loss = L_{\mathrm{d}}(z_{\mathrm{f}}, y) + \lambda_{\mathrm{R}} R(W) + \lambda_{\mathrm{p}} L_{\mathrm{p}} \tag{2.10}$$

in which λ_{p} is the regularization parameter of L_{p}.

The physics-based loss function in data-driven SHM is derived using the FE model updating outputs of the damage instances with unknown labels. For a certain damage

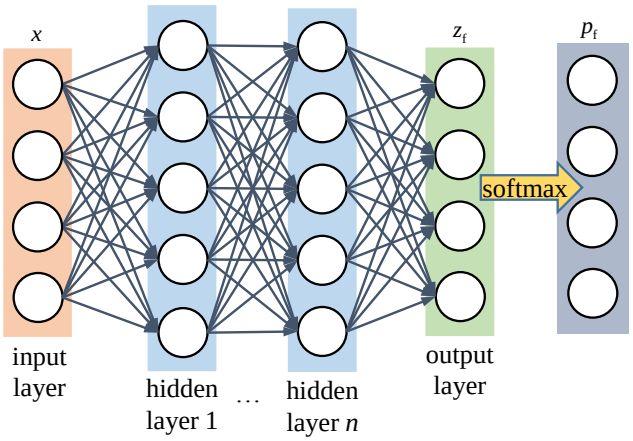


Figure 2.1. Normalizing the NN model output using the softmax function. x is the input of the NN model; z_f is the output score; p_f is the probability of each class after normalization.

case with measured modal data, FE model updating yields the damage severities at all interested locations, that is $z_{mu} = [z^1_{mu}, z^2_{mu}, ..., z^c_{mu}]$ in which c is the number of damage classes defined above. On one hand, the location with the largest damage severity can be regarded as the most probable damage location and set as the pseudo-target labels of the corresponding case, i.e., y_{mu}. Then a cross-entropy loss can be formed over all available unlabeled structural data, which is termed as L_{p1}. That is, for a certain instance x_i the loss is:

$$L_{p1}\left(z_f(x_i), y^i_{mu}\right) = -z_f^{y^i_{mu}}(x_i) + \log\left(\sum_{j=1}^{c} \exp\left(z_f^j(x_i)\right)\right) \tag{2.11}$$

On the other hand, a softmax function transforms z_{mu} to normalized damage probabilities, p_{mu} [28], that is:

$$p^i_{mu} = \frac{\exp\left(z^i_{mu}\right)}{\sum_{j=1}^{c} \exp\left(z^j_{mu}\right)} \tag{2.12}$$

for x_i. The same operation can be implemented on the output scores of NN model $z_f(x)$, which yields the predicted damage probabilities at each location, i.e., p_f as shown in Figure 2.1. Then a mean-squared-error loss can be defined between p_{mu} and p_f, which is

termed as L_{p2}. That is, for a certain instance x_i,

$$L_{\mathrm{p2}}(z_{\mathrm{f}}(x_i), z_{\mathrm{mu}}(x_i)) = \frac{1}{c}\sum_{j=1}^{c}\left(p_{\mathrm{f}}^{j}(x_i) - p_{\mathrm{mu}}^{j}(x_i)\right)^2 \tag{2.13}$$

In terms of Equations 2.8, 2.9, 2.11, and 2.13, the loss function of PGNN in Equation 2.10 can be written as:

$$\begin{aligned} Loss &= L_{\mathrm{d}} + \lambda_{\mathrm{R}} R(W) + \lambda_{\mathrm{p}} L_{\mathrm{p}} \\ &= \frac{1}{n_l}\sum_{i=1}^{n_l}\left[-z_{\mathrm{f}}^{y_i}(x_i) + \log\left(\sum_{j=1}^{c}\exp\left(z_{\mathrm{f}}^{j}(x_i)\right)\right)\right] \\ &+ \lambda_1\|W\|_1 + \lambda_2\|W\|_2 \\ &+ \lambda_{\mathrm{p1}}\frac{1}{n_u}\sum_{i=n_l+1}^{n_l+n_u}\left[-z_{\mathrm{f}}^{y_{\mathrm{mu}}^{i}}(x_i) + \log\left(\sum_{j=1}^{c}\exp\left(z_{\mathrm{f}}^{j}(x_i)\right)\right)\right] \\ &+ \lambda_{\mathrm{p2}}\frac{1}{cn_u}\sum_{i=n_l+1}^{n_l+n_u}\sum_{j=1}^{c}\left(p_{\mathrm{f}}^{j}(x_i) - p_{\mathrm{mu}}^{j}(x_i)\right)^2 \end{aligned} \tag{2.14}$$

in which n_l and n_u denote the number of available labeled and unlabeled data respectively; λ_{p1} and λ_{p2} are the regularization parameters of L_{p1} and L_{p2} respectively.

2.2.4. Implementation in PyTorch

Figure 2.2 illustrates the framework of the proposed PGML realized with an NN model, of which the details have been introduced in preceding subsections. The two dashed lines indicate physics guidance using the results of model updating on unknown damage cases, which correspond to the two physics-based loss functions L_{p1} and L_{p2} respectively.

In this study, the PGML is implemented in PyTorch, a popular deep learning

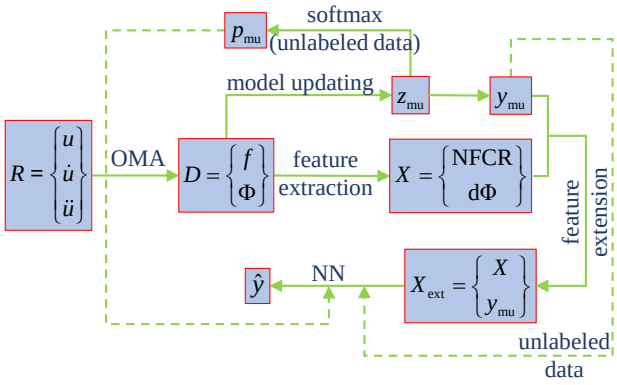


Figure 2.2. A schematic illustration of the PGML framework. R denotes the measured structural responses containing displacement u, velocity $\dot{u}$, and acceleration $\ddot{u}$; OMA stands for operational modal analysis; D is the vector containing the obtained natural frequencies f and mode shapes Φ; X denotes the feature vector extracted from D containing the normalized frequency change ratio (NFCR) and the change of mode shapes $\mathrm{d}\Phi$ of the first several modes; z_{mu} denotes the output of model updating that indicates the damage severities at each interested location; y_{mu} is the most probable damage location recommended by model updating; X_{ext} is the extended feature containing X and y_{mu}; p_{mu} contains damage probabilities of each location calculated from z_{mu} using the softmax function; $\hat{y}$ denotes the predicted labels by PGNN with the guidance of FE model updating through y_{mu} and p_{mu}.

framework based on the Torch library [29]. An MLP-NN model with designed fully-connected layers is built for each case study. ReLU [30] is selected as the activation function for each hidden layer. Adam algorithm [31] is used to perform stochastic gradient descent on the model parameters. An early stopping strategy [32] is adopted to avoid overfitting. The datasets are normalized via min-max normalization [33] before being input into the NN model. Fifty runs are implemented for each model training process with random initialization, and the model yielding the least loss is selected as the final solution.

2.3. Numerical Study

To validate the methodology of damage identification proposed in this study, this section implements a numerical case study using an FE model of a steel pedestrian bridge [24]. Figure 2.3 shows the bridge model and its sensor instrumentation. The bridge model is divided into six substructures. The elastic modulus of the frame members (E_1 to E_6) are set as the target of FE model updating in this numerical study, assuming known mass matrix and other stiffness components prior to model updating. The bridge model is instrumented with seven uniaxial and seven biaxial accelerometers that measure 21 out of the 274 DOFs. More details about the bridge model can be found in [24].

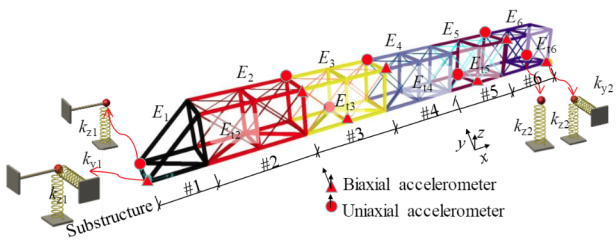


Figure 2.3. Schematic model of a steel pedestrian bridge and its sensor instrumentation [24]. The bridge structure is divided into six substructures: substructure #1 to #6. E_i (i =1,2,...,6) represents the elastic modulus of the frame members of substructure #i; E_{ti} (i =2,3,...,6) denotes the elastic modulus of the truss members of substructure #i; k_{y1} and k_{y2} are the stiffness values of transverse support springs at the two bridge ends; k_{z1} and k_{z2} are the stiffness values of vertical support springs.

Stiffness reduction is introduced into a certain substructure to simulate single-site structural damage. Multi-site damage detection is more complex using pattern recognition methods in data-driven SHM and will be studied in future research. 100 damage cases with a random damage severity between 10% to 90% are simulated for each damage location. Table 2.1 shows the damage classes and conditions simulated in this numerical

study. Modal properties of the first three modes are used to formulate the objective function in Equation 2.2 with all the weighting factors w_{λ_i} and w_{Φ_i} (i =1,2, and 3) set as 1.0. Without modeling error introduced in the established FE model, it is found in the present study that model updating yields accurate damage severities for all cases with relative error below 1%.

Table 2.1. Damage cases and conditions of the steel pedestrian bridge.

damage class	case ID	damage severity
D1	1-100	-0.1 to -0.9 on E_1
D2	101-200	-0.1 to -0.9 on E_2
D3	201-300	-0.1 to -0.9 on E_3
D4	301-400	-0.1 to -0.9 on E_4
D5	401-500	-0.1 to -0.9 on E_5
D6	501-600	-0.1 to -0.9 on E_6

However, modeling error always exists when establishing a numerical model for a real structure, which may significantly affects model updating results and the following damage evaluation quality [14]. Hence, this study focuses on damage identification considering modeling error which is modeled via increasing the support stiffness at bridge end 1 (k_{y1} and k_{z1}) by 20% and decreasing the support stiffness at bridge end 2 (k_{y2} and k_{z2}) by 20%. The mass matrix elements are also varied by changing the value of each element by a random percentage from -20% to 20% to simulate the inaccurate measurement of structural masses. Figures 2.4 and 2.5 show the results of model updating of example cases with slight (below 20%) and severe damage (above 75%) respectively. Figure 2.4 shows that the introduction of modeling error causes the model updating results significantly off

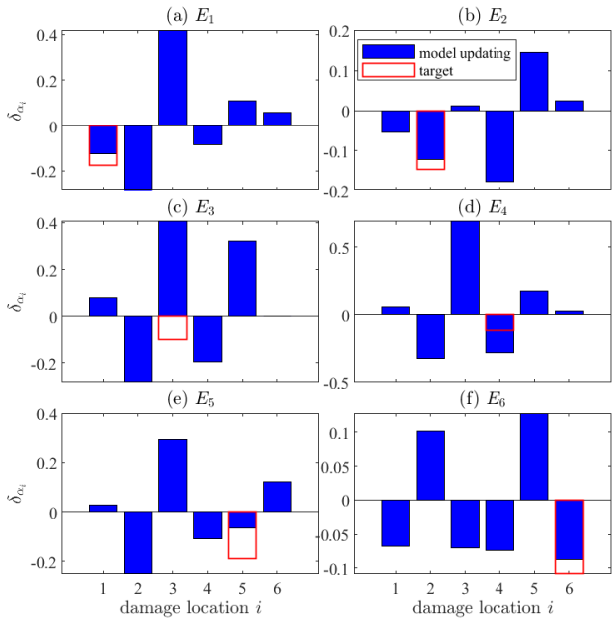


Figure 2.4. Model updating results of example cases with slight damage using the FE model with introduced modeling errors. δ_{α_i} denotes the identified damage severity at location i; E_i (i =1, 2, ..., 6) in the subfigure titles denotes the real damage location of the example case. Blue bars denote damage severities evaluated from model updating, and white bars with red edge denote target damage locations and corresponding severities.

from the correct values when slight damage happens. Incorrect damage severities are identified at target damage site and false damage is predicted at locations without damage, which makes it challenging for accurate damage quantification and localization. For example in Figure 2.4 (a), when slight damage happens to substructure 1, the identified damage severity at E_1, that is δ_{α_1}, is much smaller than that at E_2, i.e., δ_{α_2}; δ_{α_3} indicates about 40% stiffness increase, which is misleading in damage detection.Other incorrect and false damage identification results can be observed in Figures 2.4 (d) to (f).

When severe damage happens, modeling error has less adverse influence on the model updating results than when slight damage happens, as shown in Figure 2.5. The

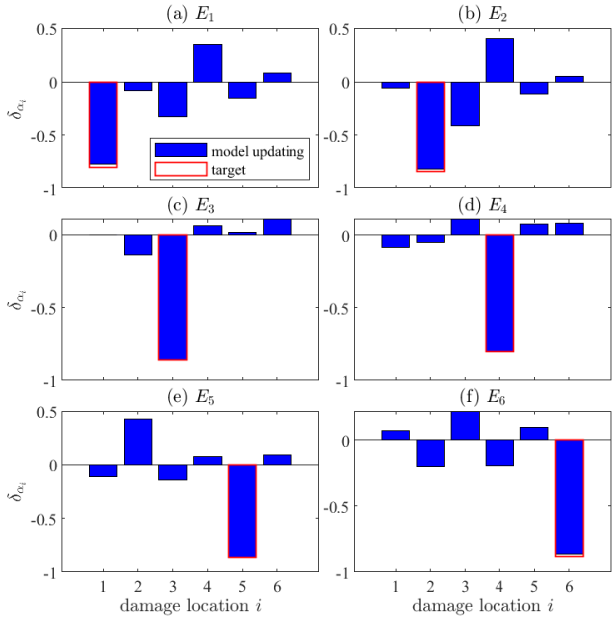


Figure 2.5. Model updating results of example cases with severe damage using the FE model with introduced modeling errors. δ_{α_i} denotes the identified damage severity at location i; E_i (i =1, 2, ..., 6) in the subfigure titles denotes the real damage locations. Blue bars denote damage severities evaluated from model updating, and white bars with red edge denote target damage locations and corresponding severities.

target damage location always has the highest damage severity that is very close to the real quantity. However, considerable misleading damage is still evaluated at locations without damage. Therefore, existence of modeling error will inevitably cause incorrect/false identification results using FE model updating. Perhaps in the case of severe damage, FE model updating can identify the damage location and quantity, yet the results are still misleading for overall damage evaluation. Moreover, in practice, accurate damage detection is necessary before severe damage happens. In a word, existence of modeling error poses a critical challenge for damage identification using deterministic FE model updating. Figure 2.6 shows the histograms of identified damage severities at correct/wrong locations obtained using model updating, in which one can observe the overall adverse effects caused

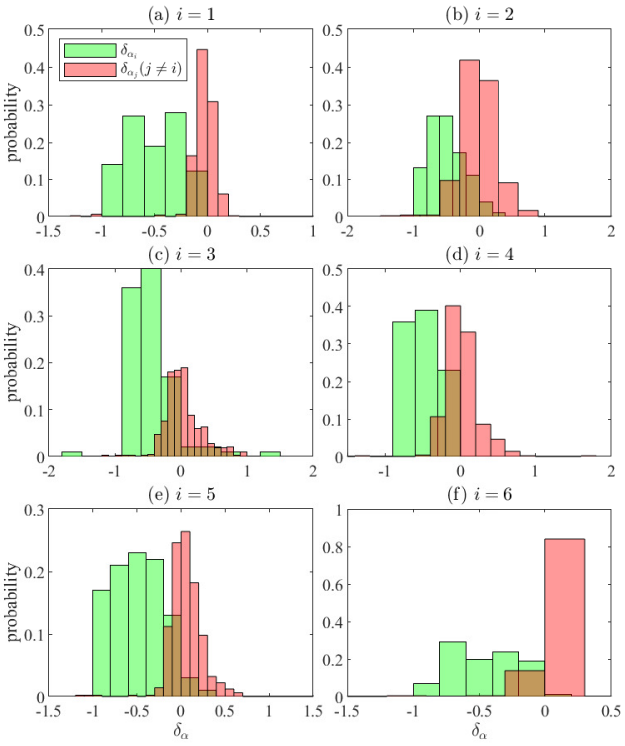


Figure 2.6. Histograms of identified damage severities from FEM mode updating. $i = 1, 2, ..., 6$ denotes the correct damage location of the plotted cases; δ_{α_i} denotes the damage severity detected at the correct location i; δ_{α_j} denotes the damage severity detected at wrong locations $j(j = 1, ..., i-1, i+1, ...6)$.

by modeling errors during model updating. Additionally, model updating yields positive damage severity at target locations for some damage cases, which can give spurious recommendations about structural conditions.

Differing from FE modeling updating that identify potential damage locations and severities simultaneously, data-driven SHM first identifies damage locations through machine learning classification; then damage severities can be evaluated through statistical regression. The present study implements damage localization through classification and evaluates the corresponding damage severities via FE model updating on the detected

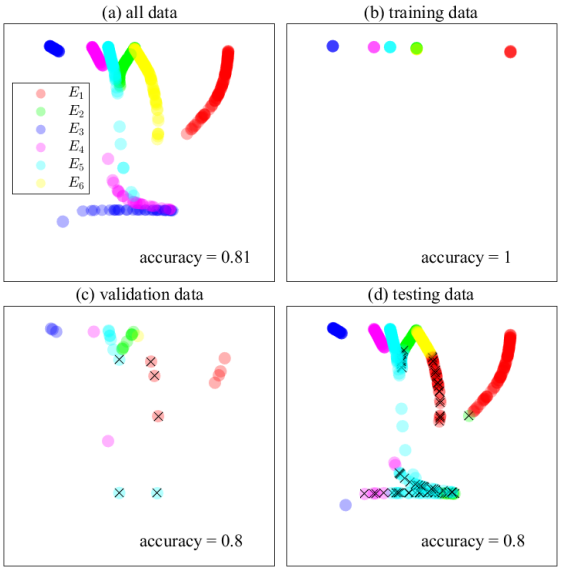


Figure 2.7. Results of damage localization on the steel pedestrian bridge model using an NN model without physics guidance. Two features are selected to visualize the data in the 2D feature space: (1) NFCR_1 (x axis); (2) $\text{d}\Phi$ of the 1^{st} measured DOF (y axis). E_i denotes the damage class corresponding to damage at location i. (a): illustration of all the data and their labels; (b)-(d): classification results on the training data, validation data, and testing data respectively; solid circles with different colors denote the predicted labels and the black cross indicates wrong labeling.

damaged sites. A three-hidden-layer MLP-NN model is designed with the number of neurons set as [50 100 50]. The data assignment is 5% for training, 5% for validation, and 90% for testing to simulate the lack of data with correct labels in reality. Additionally, slight damage is assigned to all training data, as labeled data collected in reality are usually from slight or medium damage severity scenarios. It is worth noting that this design can efficiently test the generality of the learned model in PGML in detecting future damage, while it makes pattern recognition challenging due to the lack of representative data.

For comparison, an NN model is firstly built without physics guidance by setting

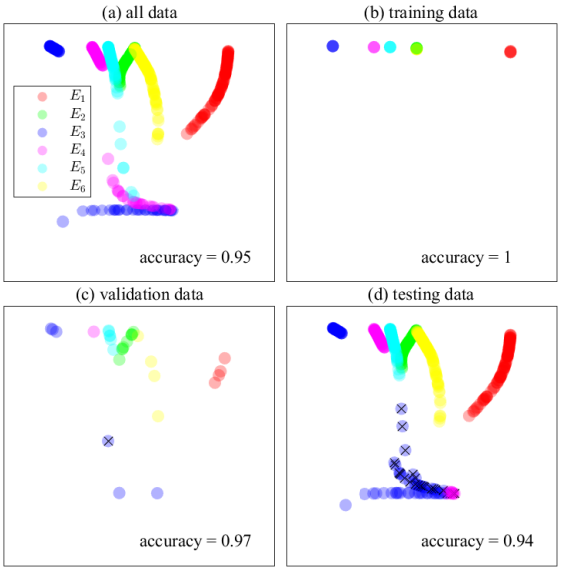


Figure 2.8. Results of damage localization on the steel pedestrian bridge model using a PGNN model with physics guidance. Two features are selected to visualize the data in the 2D feature space: (1) NFCR_1 (x axis); (2) $\text{d}\Phi$ of the 1^{st} measured DOF (y axis). E_i denotes the damage class corresponding to damage at location i. (a): illustration of all the data and their labels; (b)-(d): classification results on the training data, validation data, and testing data respectively; solid circles with different colors denote the predicted labels and the black cross indicates wrong labeling.

the corresponding regularization parameters (λ_{p1} and λ_{p2} in Equation 2.14) as zero. Figure 2.7 shows the results of damage localization on different datasets using the trained NN classifier. The data are plotted in a 2D feature space. Damage classes corresponding to the damage locations are labeled with solid circles with different colors as indicated in the legend of Figure 2.7 (a). Figure 2.7 (b) shows that all the training data are distributed in the top edge region of the feature space as a result of their limited quantity and damage severity. Hyperparameters tuning in validation fails to improve the generality of the learned model. As a result, it yields a low accuracy in model testing, as can be seen from the testing results in Figure 2.7 (d). Examining Figures 2.7 (b) through (d), one can find that correct labeling only occurs to data lying close to the training data which is very lim-

ited in representation. Hence, the generality of the learned model needs to be improved for better damage localization performance.

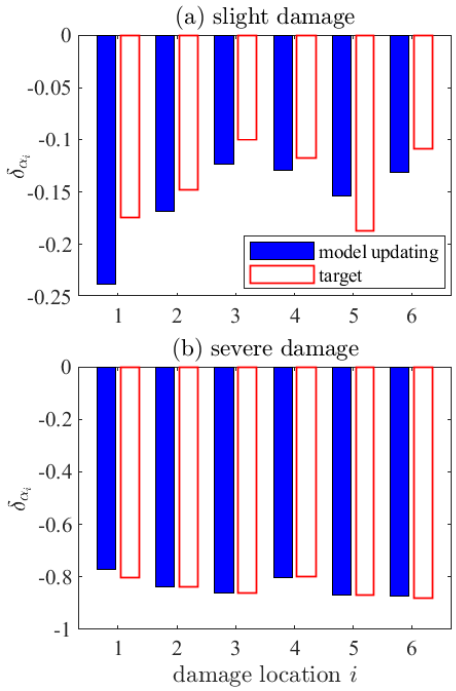


Figure 2.9. Model updating results of example cases using the FE model with introduced modeling errors with damage locations identified by the PGML model. (a) example cases with slight damage as shown in Figure 2.4; (b) example cases with severe damage as shown in Figure 2.5. Each damage case is plotted with a single bar as damage has been located. δ_{α_i} denotes the identified damage severity at location i (i =1, 2, ..., 6). Blue bars denote damage severities evaluated from model updating, and white bars with red edge denote target damage locations and corresponding severities.

To improve the generality of the learned model and its performance when tested on new cases, physics guidance is incorporated in the model learning process by activating the physics loss terms in Equation 2.14. The corresponding regularization parameters λ_{p1} and λ_{p2} are tuned using the validation dataset via a grid search strategy [34] with λ_1 and λ_2 kept unchanged from the trained NN model. Figure 2.8 shows the results of damage evaluation using the trained PGNN model. Compared with the results of NN model without

physics guidance as shown in Figure 2.7, PGNN generalizes better to the unseen testing data that has much larger quantity than either the training data or the validation data, as shown in Figure 2.8 (d). Comparing the distributions of data in the 2D feature space as shown in Figure 2.8 (b) to (d), one can find that many of the testing data are located in regions not covered by the training data or the validation data while their labels are correctly predicted by the learned classifier. This further validates the improved generality or implicit scientific consistency of the learned model with physics guidance incorporated in the model learning process.

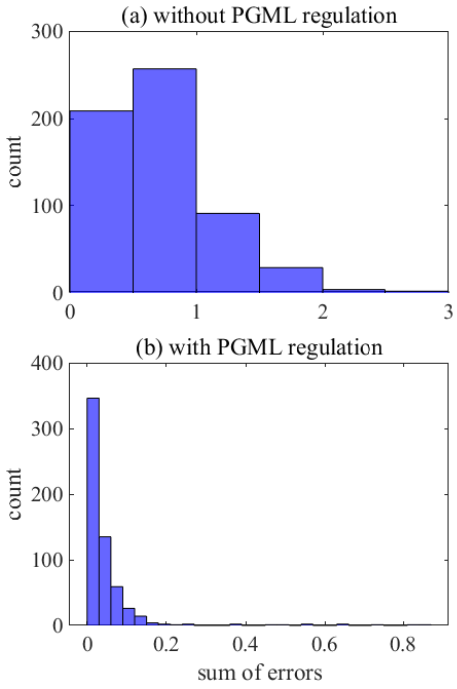


Figure 2.10. Distribution of model updating errors on the steel pedestrian bridge model without and with PGML regulation.

With most damage locations correctly identified using the learned PGNN classifier, FE model updating is rerun with the damage severity δ_{α_i} at the identified location i

set as the target of model updating. Figure 2.9 shows the results of enhanced model updating with PGML-detected damage locations. Compared with results presented in Figures 2.4 and 2.5, enhanced model updating yields damage severities closer to the target values with incorrect damage locations excluded, which further highlights the merits of PGML in structural damage evaluation when integrated with FE model updating. To further evaluate the improvement in model updating results through PGML, the error of damage severities in each case is summed over all locations. For example, the real damage severity of case 1 is $[-0.4336, 0, 0, 0, 0, 0]$, and the results of model updating without PGML is $[-0.4519, 0.0490, 0.0480, -0.0594, 0.1563, 0.1060]$, then the sum of error for case 1 is $0.0183 + 0.0490 + 0.0480 + 0.0594 + 0.1563 + 0.1060 = 0.2817$. The distribution of the sum of errors is plotted in Figure 2.10. It shows that without the regulation on damage locations from PGML, many of the error sums (22.3%) are above 1.0; however, PGML reduces the error sums of most cases significantly such that most of them (97.5%) are below 0.2. Hence, it is indicated that PGML can largely improve the damage identification performance of both data-driven and FE model updating methods..

2.4. Experimental Validation

In addition to the numerical study, an experimental study is implemented in this section using the experimental data of a 3-story frame structure published by Los Alamos National Laboratory [23]. Figure 2.11 (a) depicts the 3-story frame structure where damage was introduced by stiffness reduction of floor columns and mass addition on the 1st floor to simulate the operational and environmental influence, and bumper impacts on the

3rd floor to simulate the nonlinear behaviors of fatigue cracks. An electrodynamic shaker was used to excite the frame structure with various damage conditions with Gaussian white noise laterally on the base floor along the structural centerline. The excitation force applied from the shaker to the structure was recorded with a load cell mounted on the stringer and the structural responses were measured using four accelerometers attached at the center line of each floor as shown in Figure 2.11 (a). The data was collected and processed at a sampling frequency of 320 Hz with a data acquisition system. For each structural damage state, ten shaking tests were conducted considering the variability of excitations and structural properties. The measured data including excitations and responses can be used to extract damage-sensitive features for the structural health monitoring purposes [23].

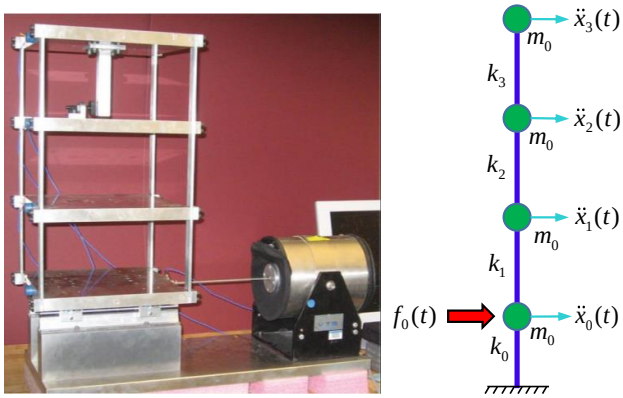


Figure 2.11. Three-story building model in [23]. (a) the LANL test structural model; (b) the schematic model. m_0 denotes the mass on each floor, k_i (i =0,1,...3) is the stiffness of the i^{th} floor; $f_0(t)$ is the excitation at the base; $\ddot{x}_i(t)$ is the acceleration measured on the i^{th} floor.

As shown in Figure 2.11 (b), this study uses a 4-DOF lumped-mass structural model to represent the 3-story building for damage detection. In FE model updating, the

mass of each floor (m_0) is calculated in terms of the plate dimension (30.5 × 30.5 × 2.5 cm^3) and the density of aluminum (2.7g/cm^3). Additionally, the sliding friction between the base floor and the rails is represented by a spring with small stiffness k_0 following the simplification by Sun and Betti [35]. As k_0 is mainly involved in the rigid body motion, its value is firstly evaluated in model updating using the 10 intact cases, which will be used subsequently for updating stiffness parameters (i.e., k_1, k_2, and k_3) with possible damage.

Table 2.2. Damage cases and conditions of the 3-story frame structure.

damage class	case ID	damage severity
D1	1-10	-0.22 on k_1
	11-20	-0.44 on k_1
D2	21-30	-0.22 on k_2
	31-40	-0.44 on k_2
D3	41-50	-0.22 on k_3
	51-60	-0.44 on k_3

The present study selects four structural conditions of the 3-story building from the database available in [23] to examine the effectiveness of PGML in damage localization. The selected conditions include the baseline condition without structural damage (termed as D0), structural condition with stiffness reduction on the first floor column(s) (D1), structural condition with stiffness reduction on the second floor column(s) (D2), and structural condition with stiffness reduction on the third floor column(s) (D3). Each damage class contains two damage severities: (1) moderate damage: 87.5% stiffness reduction of one column of a certain floor yielding 22% reduction of the floor stiffness k_i ($i = 1, 2,$ or 3); (2) severe damage: 87.5% stiffness reduction of two columns of a certain floor yielding

44% reduction of k_i. Table 2.2 lists all the damage cases used in this experimental study.

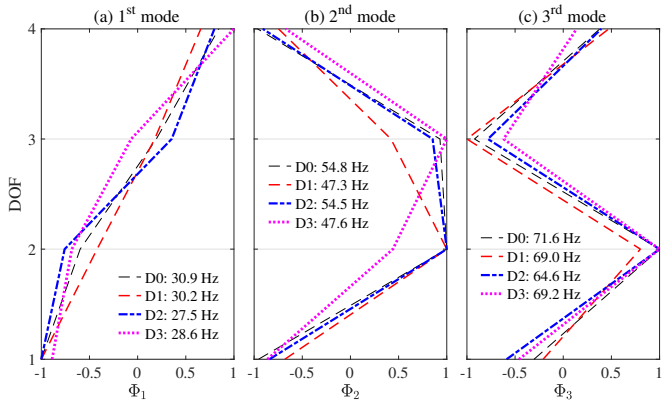


Figure 2.12. Illustration of first three mode shapes extracted from measured responses of the 3-story building. DOFs 1-4 correspond to the horizontal motions of floors. Φ_i (i =1,2, and 3) is the shape of the i^{th} mode. Di in the legend denotes the damage class defined in Table 2.2 and the value after the colon is the identified natural frequency.

The dynamic responses measured at the 4 DOFs $\ddot{x}_i(t)$ (i =0,1,...3) are processed using the frequency domain decomposition (FDD) method [36] to obtain the natural frequencies and mode shapes, which will be used to formulate the objective function in FE model updating and derive damage-sensitive features in data-driven SHM. Figure 2.12 compares the first three mode shapes of a representative case when damage happens at a certain floor. The results for case D0 are the average of modal properties obtained using the 10 intact cases. The mode shapes of each damage class presented in Figure 2.12 is obtained from an example case with severe damage (44% damage severity). The mode caused by rigid motion is excluded in this study for structural damage detection. Figure 2.12 shows that damage causes slight to obvious frequency reduction of the three modes. Obvious mode shape variations can be observed among different damage cases of each

mode. Figure 2.12 indicates that natural frequency and mode shape variations can be used to evaluate structural damage.

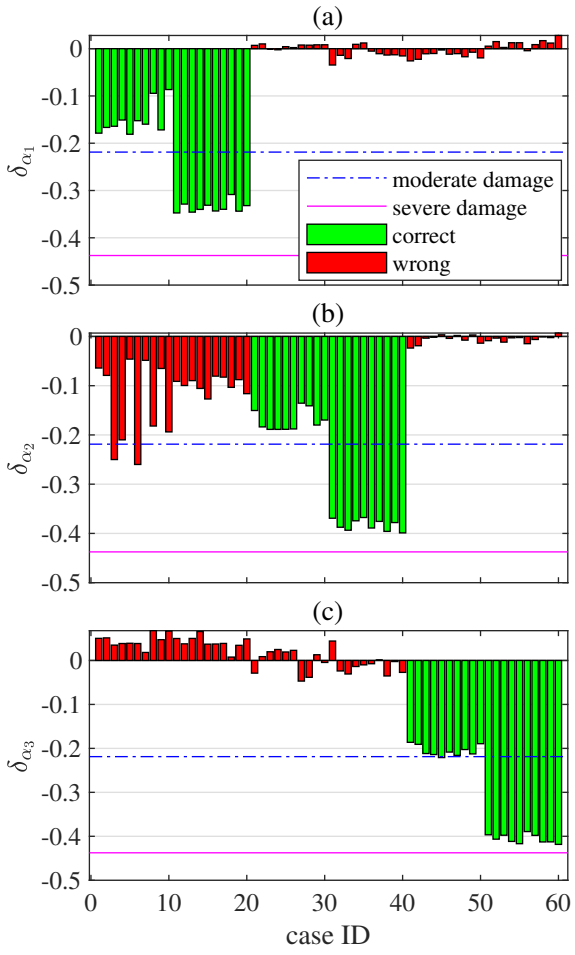


Figure 2.13. Model updating based damage identification results of the 3-story building using the 4-DOF numerical model. δ_{α_i} denotes the damage severities on k_i (i =1,2, and 3) evaluated through comparing the updated stiffness values with that of the intact structure. Green bars labeled as "correct" in the legend indicate evaluated damage at correct locations, and red bars denote identified false damage at locations without damage. The two reference lines indicate the damage severities designed in the experimental tests, that is -0.22 for moderate damage and -0.44 for severe damage. See Table 2.2 for more details about the damage cases.

With the experimental data, FE model updating is implemented using the extracted modal properties to demonstrate the limitation of model updating in presence of

modeling error. Figure 2.13 shows the identified damage results using model updating and the measured modal properties. It shows that model updating can detect damage occurrence in the 60 cases as indicated by the "correct" bars. However, notable errors exist in the damage evaluation results due to the considerable modeling error in establishing the numerical model (e.g., mass measurement and connection modeling), as shown by the "wrong" bars; especially when moderate damage happens on k_1 (case 1-10), significant wrong damage is identified on k_2, which may cause error in damage localization. It should be noted that the experimental tests in LANL contain no slight damage case, for which model updating results might be influenced by significant modeling error, as indicated in the numerical case study.

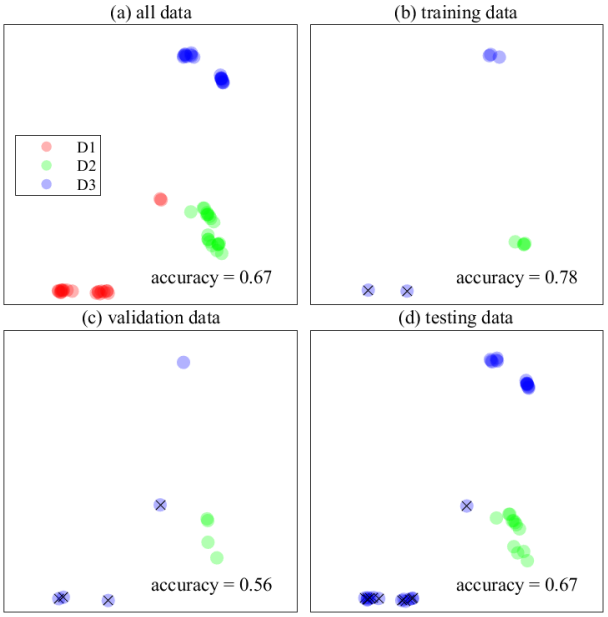


Figure 2.14. Results of damage localization on the 3-story building using an NN model without physics guidance. The x axis is PC1 of the data, and the y axis is PC2. Di (i =1,2,and 3) denotes the damage cases as listed in Table 2.2. Solid circles with different colors denote the predicted labels and the black cross indicates wrong labeling. The damage detection accuracy of each dataset is shown in the subfigures.

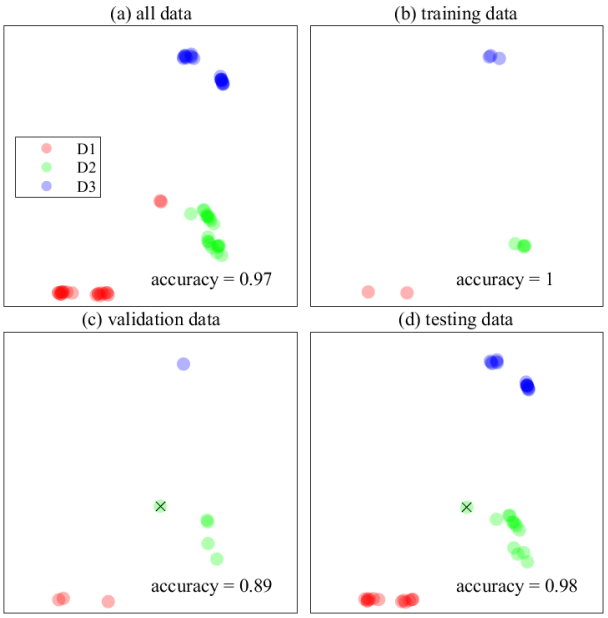


Figure 2.15. Results of damage localization on the 3-story building using a PGNN model with physics guidance. The x axis is PC1 of the data, and the y axis is PC2. Di (i =1,2,and 3) denotes the damage cases listed in Table 2.2. Solid circles with different colors denote the predicted labels and the black cross indicates wrong labeling. The damage detection accuracy of each dataset is shown in the subfigures.

The same NN structure as used in the numerical study section is used in this experimental study with all the hyperparameters retuned. The data assignment is 15% for training, 15% for validation, and 70% for testing, so that each dataset contains data from different damage classes. Figure 2.14 shows the results of classification for damage detection. The data are plotted using the two principal components (PCs) of the features obtained from principal component analysis (PCA). Figure 2.14 indicates that with the hyperparameters (λ_1 and λ_2 in Equation 2.14) tuned in validation, the accuracy on the training data is 0.78. Without physics guidance, the trained NN model yields an accuracy as low as 0.67 on the testing data because of insufficient data with correct labels for learning a model with satisfactory generality.

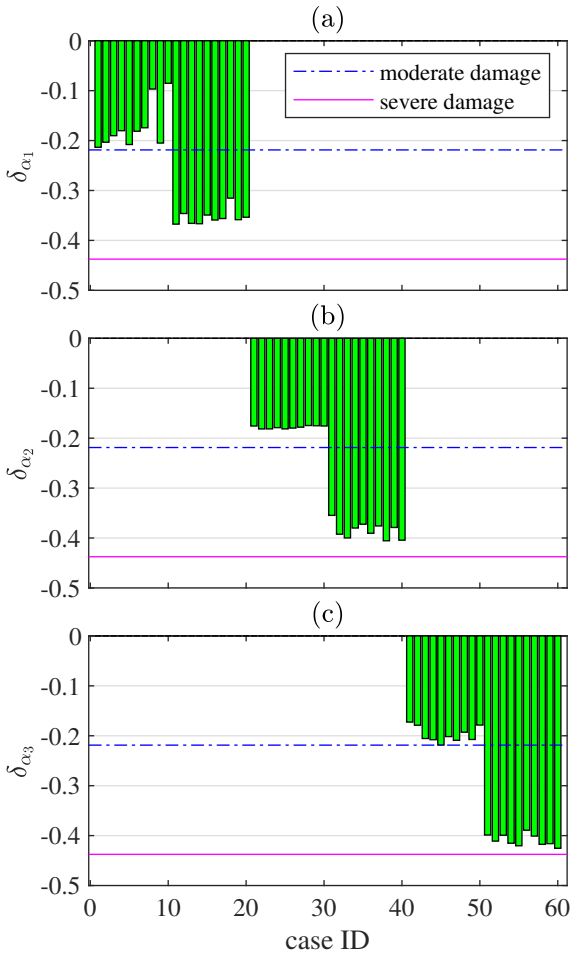


Figure 2.16. Model updating results of the 3-story building using the 4-DOF numerical model with damage location identified by PGML. δ_{α_i} denotes the damage severities on k_i (i =1,2, and 3) evaluated through comparing with the updated stiffness values of the intact structure. The two reference lines indicate the damage severities designed in the experimental tests, that is -0.22 for moderate damage and -0.44 for severe damage. See Table 2.2 for more details about the damage cases.

Figure 2.15 shows the results of damage localization with a PGNN model, in which the regularization parameters for the physics loss (λ_{p1} and λ_{p2}) are tuned in validation with λ_1 and λ_2 kept unchanged from the tuned NN model above. Comparison between Figure 2.14 and Figure 2.15 shows that PGNN improves the damage detection accuracy of all datasets through incorporating physics guidance obtained in model updating in the

classifier's learning process. It should be noted that although the correct labels of the validation data is not contained in the training process, and that the labels evaluated from model updating (y_{mu} in Figure 2.2) contain errors due to some of the misleading model updating results, the PGNN can provide consistently accurate damage identification results.

Figure 2.16 shows the results of model updating with damage localized via PGML. With incorrect damage locations excluded in model updating, the identified damage severities at correct sites generally get closer to the correct values as indicated by the reference lines. The distribution of the sum of errors is plotted in Figure 2.17. It shows that without the regulation on damage locations from PGML, many of the error sums (20%) are above 0.2; however, PGML reduces the error sums of most cases significantly such that most of them (93.3%) are below 0.1. Hence, it can be concluded that the PGML can largely improve the damage identification performance.

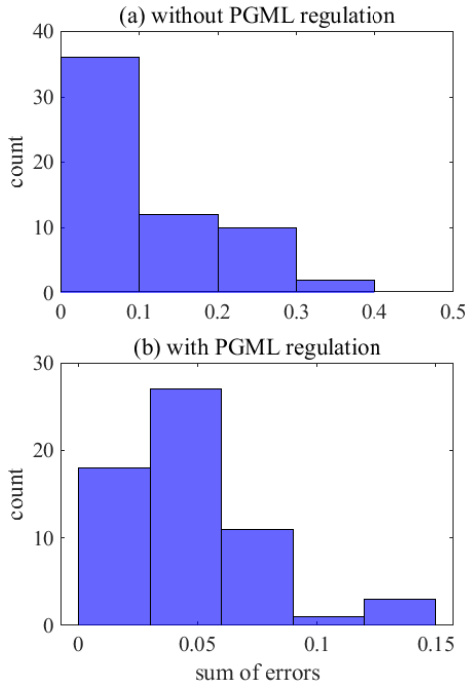


Figure 2.17. Distribution of model updating errors on the 3-story building with and without PGML regulation.

2.5. Summary

This study proposes a physics-guided machine learning (PGML) method to integrate the data-driven SHM with FE model updating, so that the merits of the two categories of SHM approaches can be preserved and the negative influence of their shortcomings can be mitigated. Regarding an established FE model and its updated outputs as an implicit representation of underlying physics of the target structure, this study implements PGML through incorporating the results of FE model updating into a neural network (NN) model to have a physics-guided neural network (PGNN) model. On one hand, the original feature vector is extended with output labels predicted from FE model updating; on the other hand, the outputs of FE model updating, including the output scores and output labels, are both incorporated into the objective function of the modified NN

model. These modifications of the NN model with respect to its inputs and objective function formulation are expected to improve the physics-related consistency of the learned model. Additionally, with the most probable damage location identified from PGML, FE model updating can be rerun with constrained targets, which may largely improve the efficiency of solving the optimization problem and increase its accuracy. Hence, the adverse effects of modeling error can be significantly mitigated.

A numerical case study is implemented with an FE model of a steel pedestrian bridge where the elastic modulus of the frame members are set as the target of FE model updating and damage identification. Features are extracted from modal properties following previous studies in literature. It shows that before implementing PGML, FE model updating yields misleading values of target parameters due to significant modeling error when establishing the FE model, and an NN model learned with the limited data cannot generalize well when applied on new testing data. In comparison, PGML can not only improve the generality of the learned NN model with testing accuracy increased from 0.8 to 0.94 but enhance the performance of FE model updating in damage detection with largely reduced sum of errors. These improved damage identification outcome validates the effectiveness of PGML in integrating data-driven and model-based SHM for better performance. An experimental study with a 3-story frame structure further validates the effectiveness of PGML in structural damage identification with either data-driven or FE model updating method. Therefore, it can be concluded that PGML has the potential of mitigating the challenge of data insufficiency in data-driven SHM and the potential of reducing the effects of modeling error in FE model updating. However, it is noted by the author that FE model updating outputs may not be the best representation of scientific rules for

structural damage evaluation, which will be one of the focuses of future research.

Chapter 3. Transfer-Learning Guided Bayesian Model Updating for Damage Identification Considering Modeling Uncertainty

3.1. Introduction

Chapter 2 introduces physics guidance from FE model updating into the pattern recognition process in data-driven damage identification, which improves the scientific consistency of the damage classification results with the learned machine learning model. The present chapter presents another methodology to integrate the data-driven and model-based methods through domain adaptation.

Finite element model updating using vibration based features (e.g., modal properties) has been widely used for structural identification and damage detection [37, 38, 39]. It has the potential of identifying the structural damage condition from the variation of identified structural parameters [40]. Finite element model updating methods fall into two main categories: deterministic methods and probabilistic methods. As pointed out in the literature [3, 41, 15, 42, 43], model updating needs to account for three types of uncertainties when applied to real-world structures : (1) measurement noise and feature identification error; (2) inherent parameter variations caused by ambient and/or environmental factors; (3) modeling uncertainty or modeling error due to model simplification and omission. With the first two types of uncertainties comprehensively considered in model updating [41, 44, 45, 46, 47, 48], the third type of uncertainty is widely recognized as the most challenging and critical in model updating [39, 41, 49, 50].

Modeling uncertainty is caused by the simplifications applied to the physical problem in finite element analysis [51]. Differing from the measurement uncertainty that typ-

ically has a zero mean and can be largely reduced by averaging the collected data sets [13], modeling uncertainty is a bias error and causes a shift in the structural responses and measured features which are usually not centered about zero. Major sources of modeling uncertainty can be categorized into the following two types [37, 52, 49, 53, 54, 13]: (1) model parameter uncertainty in measuring or quantifying important structural parameters, such as material properties, geometry of structural components, properties of nonstructural components, etc; (2) modeling principle uncertainty from simplifying assumptions or lack of understanding of the real structural system in simulating the structural characteristics, such as boundary conditions, material constitutive law, damping simulation in dynamics, governing equations, model order, discretization in FE models, etc. Modeling uncertainty is recognized as one important barrier limiting the in-practice implementation of model updating in structural health monitoring, because it inherently reduces the capability of structural identification to capture the real physics underlying a structure [13]. In model-based SHM, modeling uncertainty causes significant challenge in structural parameter estimation [13, 55], structural damage detection [14, 56, 57] and predicting structural features and responses [54, 15, 55, 39, 47].

Using the sensitivity-based deterministic model updating, Sanayei et al. [13] compared the performance of four error functions in the presence of modeling uncertainty in terms of propagating the influence of modeling uncertainty and the accuracy of final parameter estimates, and they revealed the superiority of stiffness-based error functions over the flexibility-based error functions for both static and modal formulations. Goulet and Smith [43, 58] used the error-domain model falsification (EDMF) technique to provide candidate model sets in the presence of systematic errors. It is shown that the EDMF tech-

nique can falsify model instances/classes for compatibilities between estimates and measurements. However, the computational cost can be demanding and the damage identification accuracy will be compromised in the presence of modeling uncertainty.

On the other hand, the prediction error between the model output and the measurements can be used as a means of considering effects of modeling uncertainty in model updating [59]. The distribution of the prediction error is usually assumed as Gaussian with a non-zero mean based on the principle of maximum entropy [54, 60]. To account for prediction error in probabilistic model updating, Behmanesh et al. [15, 61] and Song et al. [3, 41] used a hierarchical Bayesian model updating method to update the mean $\boldsymbol{\mu}_{\mathbf{e}}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{e}}$ of the prediction error (denoted as $\mathbf{e}$) assuming that $\mathbf{e}$ in a certain test $\mathbf{e}_t$ follows a Gaussian distribution, that is $\mathbf{e}_t \sim N(\boldsymbol{\mu}_{\mathbf{e}}, \boldsymbol{\Sigma}_{\mathbf{e}})$. Then the updated $\boldsymbol{\mu}_{\mathbf{e}}$ and $\boldsymbol{\Sigma}_{\mathbf{e}}$ can be used to quantify the uncertainty caused by modeling error. However, as mentioned in References [15, 55, 39], this method using hierarchical Bayesian model updating cannot accurately estimate the values of structural parameters because of the compensation effects in the presence of modeling uncertainty.

Based on the assumption that combining multiple model predictions yields an improved prediction accuracy, References [62, 63, 64] used Bayesian model averaging (BMA) and BMA with an adjustment factor [65, 66] to quantify the model selection uncertainty [67]. However, the effectiveness of these methods is based on the premise that the true/best model is among the set of models considered, which can't be guaranteed in real engineering. Finding that modeling uncertainty has larger influence on higher vibrational modes, Behmanesh et al. [49] attempted to reduce the effects of modeling uncertainty by selecting the optimal subsets of modal data through model selection and

BMA. The authors proposed a new likelihood function using data sets under damaged and intact conditions. However, the proposed error functions and the corresponding likelihood function cannot inherently solve the problem of model updating caused by modeling uncertainty, because the prediction error still has a non-zero mean and thus the updated structural parameters remain biased from the true values. Moreover, this method remains to be validated in case of remarkable modeling uncertainty and/or severe structural damage. In summary, while extensive efforts have been devoted to analyzing and reducing the influence of modeling uncertainty in model updating, it remains a challenging problem in model-based SHM. This study aims to reduce the influence of modeling error in model-based SHM through guiding the model updating process with a domain-adaptation based pattern recognition method.

In structural identification, modeling error exists and renders the established and updated numerical model intrinsically deviant from the target real structure. As a result, the feature data from the numerical model and physical structure lie in two disparate domains, namely, the source domain and the target domain, respectively. In contrast to data measured from the target structure that are usually sparse, noisy, and incomplete, there is almost no limitation to the quality and quantity of the data that can be obtained from the source structures. Hence, one can develop a model (e.g., a machine learning model) for structural condition identification using sufficient data from the source structure. Domain adaptation is designed for solving a machine learning problem in the target domain by taking advantage of the training data and/or the learned knowledge in the source domain that differs significantly from while keeps similar to the target domain [68]. It can potentially improve the learning performance in case the learning task lies in one domain

having insufficient labeled training data but sufficient data is available in another domain [69]. Domain adaptation can be realized through transfer learning which generally transfers data instances, features, or learned models across different but related tasks or domains [69, 70]. Liu and Worden proposed for the first time applying domain adaptation on structural health monitoring [71]. Model-based SHM in the presence of significant modeling error fits well into the scenario of domain adaptation in the following aspects: (1) modeling error produces significant divergence between these two though related domains; (2) a lot more information can be obtained from the source structure than from the target structure.

Moreover, model based SHM and data based SHM have been extensively studied in literature and widely applied in practice [72, 73, 4, 10]. Each type of methods has its merits as well as shortcomings. In model based SHM, FE model establishment and updating require onerous model tuning and are sussceptible to remarkable uncertainties, which can be largely mitigated in the model-free data-driven methods [11]. On the other hand, model based methods provide an updated physics-based numerical model that can be used for damage prognosis. Therefore, it would be promising to combine these two types of SHM methods to strengthen their merits and mitigate the shortcomings [74].

To this end, this study proposes using the domain adaptation technique via transfer learning (TL) to improve the performance of model updating for structural identification and damage detection in the presence of significant modeling uncertainty. TL is realized through adaptation regularization via minimizing the distribution difference and manifold discrepancy between the source and target domains in addition to reducing the prediction error and model complexity as done in traditional machine learning. With the guidance

from TL regarding damage locations, Bayesian model updating (BMU) is implemented to estimate structural parameters and their uncertainties as well as the most probable damage severities. The results of numerical and experimental studies indicate that this TL-guided BMU (BMU-TL) has the potential of reducing the influence of model uncertainty in model updating and improving the overall performance of damage identification.

3.2. Framework of Hierarchical Bayesian Model Updating for Probabilistic Structural Identification

This section elaborates the framework of hierarchical Bayesian model updating used in the present study for probabilistic structural identification and damage detection, which covers the following modules: (1) selection of structural model class; (2) selection of the prior distributions, error function, and likelihood function; (3) derivation of the posterior distribution and full conditional distributions; (4) approximate maximum a posteriori (MAP) estimation of parameters. This framework will be used in the rest of this chapter to compare the performance of BMU and BMU-TL in structural identification and damage identification.

3.2.1. Structural Model Class

In linear structural identification, the model class $\mathcal{M}^0$ assumes a known mass matrix $\boldsymbol{M}$, and the stiffness matrix $\boldsymbol{K}$ can be expressed as a linear combination of $N_\mathrm{p}+1$ substructural stiffness matrices $\boldsymbol{K}_p$ with $p=0,1,...,N_\mathrm{p}$, that is:

$$\boldsymbol{K}=\boldsymbol{K}_0+\sum_{p=1}^{N_\mathrm{p}}\theta_p\boldsymbol{K}_p \tag{3.1}$$

in which $\boldsymbol{K}_0$ is the known substructural stiffness without uncertainty, $\boldsymbol{K}_p$ is the nominal stiffness contribution of the p^{th} substructure to the global stiffness matrix $\boldsymbol{K}$, and θ_p is the

p^{th} scaling parameter to be estimated though model updating using the measured data. Then the target of model updating becomes $\boldsymbol{\theta} = \left\{\theta_1, \theta_2, ...\theta_{N_\text{p}}\right\}$ with N_p being its dimension.

In vibration-based structural identification, the measured data $\widetilde{\mathcal{D}}$ for model updating usually includes the identified eigenfrequencies $\widetilde{\boldsymbol{\lambda}}$ and the corresponding mode shapes $\widetilde{\boldsymbol{\Phi}}$, that is $\widetilde{\mathcal{D}} = \left\{\widetilde{\boldsymbol{\lambda}}, \widetilde{\boldsymbol{\Phi}}\right\}$. The symbol with a tilde (e.g. $\widetilde{\mathcal{D}}$) denotes the measurement from experimental tests, which is a counterpart of the estimate from the established or updated model using the supposed values of the parameter set $\boldsymbol{\theta}$ (e.g. $\mathcal{D}(\boldsymbol{\theta})$). Suppose that N_m modes of structural vibrations can be identified from modal testing with sufficient accuracy, then $\widetilde{\boldsymbol{\lambda}}$ and $\widetilde{\boldsymbol{\Phi}}$ can be expressed as follows:

$$\widetilde{\boldsymbol{\lambda}} = \left\{\widetilde{\lambda}_1, \widetilde{\lambda}_2, ..., \widetilde{\lambda}_{N_\text{m}}\right\} \quad \text{and} \quad \widetilde{\boldsymbol{\Phi}} = \left\{\widetilde{\boldsymbol{\Phi}}_1, \widetilde{\boldsymbol{\Phi}}_2, ..., \widetilde{\boldsymbol{\Phi}}_{N_\text{m}}\right\} \tag{3.2}$$

in which $\widetilde{\boldsymbol{\Phi}}_m$ ($m = 1, 2, ..., N_\text{m}$) is a vector with a length equal to the number of measured DOFs with pre-installed sensors (i.e., N_s). With the structural model class selected and the model updating problem properly defined, the Bayesian inference method can be used for probabilistic structural parameter estimation.

3.2.2. Priors, Likelihood Function, Posterior Distribution, and Full Conditional Distributions

In reality, more than one tests can be conducted for structural parameter estimation that may correspond to different structural conditions due to ambient/environmental uncertainties. Hence, following the principle of maximum entropy, this study assumes that the parameter vector in the t^{th} test $\boldsymbol{\theta}_t$ ($t = 1, 2, ...N_t$) follows a truncated multivariate

Gaussian distribution ($\theta_p > 0$, $p = 1, 2, ..., N_\mathrm{p}$), that is:

$$\boldsymbol{\theta}_t \sim \mathcal{N}^+ \left(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta \right) \tag{3.3}$$

in which the superscript $^+$ denotes a truncated multivariate Gaussian distribution in the positive space.

Additionally, the parameters $\theta_1, \theta_2, ..., \theta_{N_\mathrm{p}}$ in each test are assumed independent from each other, which renders the covariance matrix $\boldsymbol{\Sigma}_\theta$ diagonal. That is:

$$\boldsymbol{\Sigma}_\theta = \begin{bmatrix} \sigma_{\theta,1}^2 & & & \\ & \sigma_{\theta,2}^2 & & \\ & & \ddots & \\ & & & \sigma_{\theta,N_\mathrm{p}}^2 \end{bmatrix} \tag{3.4}$$

The error function (i.e., the prediction error) of the t^{th} test is defined as the discrepancy between the model-calculated modal parameters using the parameter set (i.e., $\boldsymbol{\theta}_t$) and that obtained from their measured counterparts. The error terms from the eigenfrequency ($\boldsymbol{e}_{\lambda_t}$) and mode shape ($\boldsymbol{e}_{\boldsymbol{\Phi}_t}$) are respectively defined as follows:

$$\boldsymbol{e}_{\lambda_{mt}} = \frac{\tilde{\lambda}_{mt} - \lambda_m(\boldsymbol{\theta}_t)}{\tilde{\lambda}_{mt}} \quad \text{and} \quad \boldsymbol{e}_{\boldsymbol{\Phi}_{mt}} = \frac{\widetilde{\boldsymbol{\Phi}}_{mt}}{\|\widetilde{\boldsymbol{\Phi}}_{mt}\|} - a_{mt} \frac{\boldsymbol{\Gamma}\boldsymbol{\Phi}_m(\boldsymbol{\theta}_t)}{\|\boldsymbol{\Gamma}\boldsymbol{\Phi}_m(\boldsymbol{\theta}_t)\|} \tag{3.5}$$

in which $\lambda_m(\boldsymbol{\theta}_t)$ and $\boldsymbol{\Phi}_m(\boldsymbol{\theta}_t)$ are the evaluated m^{th} eigenfrequency and mode shape from the model using the parameter values $\boldsymbol{\theta}_t$; $\tilde{\lambda}_{mt}$ and $\widetilde{\boldsymbol{\Phi}}_{mt}$ are the measured counterparts; $\boldsymbol{\Gamma}$ is the mapping matrix that converts the model-calculated mode shapes to the measured DOFs; a_{mt} is the cosine of the angle between the measured and model-calculated mode shapes after conversion, which is defined as:

$$a_{mt} = \frac{\widetilde{\boldsymbol{\Phi}}_{mt}^{\mathrm{T}} \boldsymbol{\Gamma}\boldsymbol{\Phi}_m(\boldsymbol{\theta}_t)}{\|\widetilde{\boldsymbol{\Phi}}_{mt}\| \|\boldsymbol{\Gamma}\boldsymbol{\Phi}_m(\boldsymbol{\theta}_t)\|} \tag{3.6}$$

in which the superscript $^{\mathrm{T}}$ denotes the transposed matrix.

Similar to the parameter set θ_t, this study assumes that the error function in each test follows a Gaussian distribution. That is:

$$\boldsymbol{e}_t = \begin{bmatrix} \boldsymbol{e}_{\lambda_t} \\ \\ \boldsymbol{e}_{\boldsymbol{\Phi}_t} \end{bmatrix} \sim \mathcal{N}\left\{\boldsymbol{\mu}_{\mathrm{e}}, \boldsymbol{\Sigma}_{\mathrm{e}}\right\} \tag{3.7}$$

which yields the expression of the likelihood function of measured data $\widetilde{\mathcal{D}}$. That is, given the value of $\boldsymbol{\theta}_t$, the difference of $\widetilde{\mathcal{D}}_t$ from its predicted counterpart $\mathcal{D}_t$ follows the Gaussian distribution given above in Equation 3.7.

It should be noted that the error mean $\boldsymbol{\mu}_{\mathrm{e}}$ is not set as zero considering the modeling bias that causes a shift of the model-calculations from the truth and the measurements. The covariance matrix $\boldsymbol{\Sigma}_{\mathrm{e}}$ is set as diagonal assuming independence between error terms, that is:

$$\boldsymbol{\Sigma}_{\mathrm{e}} = \begin{bmatrix} \sigma^2_{\mathbf{e},1} & & & \\ & \sigma^2_{\mathbf{e},2} & & \\ & & \ddots & \\ & & & \sigma^2_{\mathbf{e},N_{\mathrm{e}}} \end{bmatrix} \tag{3.8}$$

in which N_{e} is the number of error terms and $N_{\mathrm{e}} = N_{\mathrm{m}}(N_{\mathrm{s}} + 1)$.

Regarding the hyper-priors (including that of $\boldsymbol{\mu}_\theta$, $\boldsymbol{\mu}_{\mathrm{e}}$, $\boldsymbol{\Sigma}_\theta$, and $\boldsymbol{\Sigma}_{\mathrm{e}}$) with hyperparameters, this study adopts a truncated uniform prior for $\boldsymbol{\mu}_\theta$, a non-informative uniform prior for $\boldsymbol{\mu}_{\mathrm{e}}$, and non-informative inverse-gamma priors for the diagonal terms of covariance matrices σ^2_p and $\sigma^2_{\mathbf{e},p_{\mathrm{e}}}$, that is:

$$\mu_{\theta_p} \sim U(0, \mu^{\mathrm{u}}_{\theta_p}) \tag{3.9}$$

$$p\left(\boldsymbol{\mu}_{\mathrm{e}}\right) \propto 1 \tag{3.10}$$

$$\sigma^2_{\theta,p} \sim InvGamma(\alpha_\theta, \beta_\theta) \tag{3.11}$$

$$\sigma^2_{\mathbf{e},p_\mathrm{e}} \sim InvGamma(\alpha_\mathrm{e}, \beta_\mathrm{e}) \tag{3.12}$$

in which $\boldsymbol{\mu}^{\mathrm{u}}_\theta$ is the upperlimit of the structural parameter mean that can be determined by conservatively evaluating the structural parameters from engineering drawings or surveying the values of similar structures, $\alpha_\theta = \alpha_\mathrm{e} = 1$ are the shape parameters of the inverse-gamma distribution, and $\beta_\theta = \beta_\mathrm{e} = 2$ are their scale parameters.

Figure 3.1 illustrates the graphical model of hierarchical Bayesian inference for structural model updating that summarizes the priors and likelihood function explained above and their relationships, which is also beneficial for deriving the posterior distribution and full conditional distributions. Based on the Bayes' theorem, the posterior probability density function (PDF) of the parameter set is proportional to the likelihood function multiplied by the prior PDFs. That is:

$$\begin{aligned} &p\left(\boldsymbol{\Theta}, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \boldsymbol{\mu}_\mathrm{e}, \boldsymbol{\Sigma}_\mathrm{e} \middle| \widetilde{\boldsymbol{\lambda}}, \widetilde{\boldsymbol{\Phi}}\right) \propto \\ &\prod_{t=1}^{N_t} p\left(\widetilde{\boldsymbol{\lambda}}, \widetilde{\boldsymbol{\Phi}} \middle| \boldsymbol{\theta}_t, \boldsymbol{\mu}_\mathrm{e}, \boldsymbol{\Sigma}_\mathrm{e}\right) p\left(\boldsymbol{\theta}_t \middle| \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta\right) p\left(\boldsymbol{\mu}_\theta\right) p\left(\boldsymbol{\Sigma}_\theta\right) p\left(\boldsymbol{\mu}_\mathrm{e}\right) p\left(\boldsymbol{\Sigma}_\mathrm{e}\right) \end{aligned} \tag{3.13}$$

in which $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_{N_t}\}$. The first part with the product sign is the likelihood function that originates from the distribution of the error function defined in Equation 3.7.

Plugging in the expressions of priors and likelihood function, the posterior PDF

becomes

$$
\begin{aligned}
&p\left(\boldsymbol{\Theta}, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \boldsymbol{\mu}_\mathrm{e}, \boldsymbol{\Sigma}_\mathrm{e} \middle| \widetilde{\boldsymbol{\lambda}}, \widetilde{\boldsymbol{\Phi}}\right) \propto \\
&\left[\prod_{p_\mathrm{e}=1}^{N_\mathrm{e}} \left(\sigma_{\mathrm{e},p_\mathrm{e}}^2\right)^{-\frac{N_t}{2}}\right] \times \exp\left(-\sum_{t=1}^{N_t} J_{\mathrm{e},t}\right) \times \left[\prod_{p=1}^{N_\mathrm{p}} \left(\sigma_{\theta,p}^2\right)^{-\frac{N_t}{2}}\right] \times \exp\left(-\sum_{t=1}^{N_t} J_{\theta,t}\right) \times \\
&\left[\prod_{p=1}^{N_\mathrm{p}} \left(\sigma_{\theta,p}^2\right)^{-\alpha_\theta-1} \exp\left(-\frac{\beta_\theta}{\sigma_{\theta,p}^2}\right)\right] \times \left[\prod_{p_\mathrm{e}=1}^{N_\mathrm{e}} \left(\sigma_{\mathrm{e},p_\mathrm{e}}^2\right)^{-\alpha_\mathrm{e}-1} \exp\left(-\frac{\beta_\mathrm{e}}{\sigma_{\mathrm{e},p_\mathrm{e}}^2}\right)\right] \times \\
&\prod_{p=1}^{N_\mathrm{p}} 1\left(0 < \mu_{\theta_p} < \mu_{\theta_p}^\mathrm{u}\right)
\end{aligned}
\tag{3.14}
$$

in which $1(A)$ is the indicator function of the event A; $J_{\mathrm{e},t}$ and $J_{\theta,t}$ are defined as:

$$
J_{\mathrm{e},t} = \frac{1}{2}\left(\boldsymbol{e}_t - \boldsymbol{\mu}_\mathrm{e}\right)^\mathrm{T} \left(\Sigma_\mathrm{e}\right)^{-1} \left(\boldsymbol{e}_t - \boldsymbol{\mu}_\mathrm{e}\right) \tag{3.15}
$$

$$
J_{\theta,t} = \frac{1}{2}\left(\boldsymbol{\theta}_t - \boldsymbol{\mu}_\theta\right)^\mathrm{T} \left(\Sigma_\theta\right)^{-1} \left(\boldsymbol{\theta}_t - \boldsymbol{\mu}_\theta\right) \tag{3.16}
$$

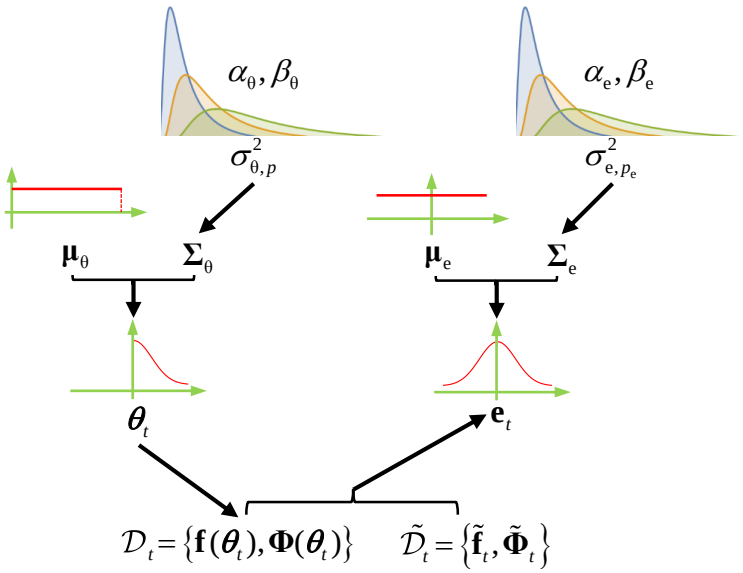


Figure 3.1. Graphical model for hierarchical Bayesian model updating.

The posterior full conditional distribution of a parameter (or a parameter set) is the conditional distribution of that parameter given current values of all other parameters.

For a certain parameter v, the rest of parameters can be denoted as V_{-v}, then the full conditional distribution $P(v|V_{-v})$ has the form [75]:

$$\begin{aligned} P(v|V_{-v}) &\propto P(v, V_{-v}) \\ &= P(v|\text{parents}[v]) \times \prod_{w \in \text{children}[v]} P(w|\text{parents}[w]) \end{aligned} \tag{3.17}$$

in which parents of a parameter are the parameters with an arrow pointing to v and children of a parameter denote parameters on a directed path starting from it, as shown in Figure 3.1. Following this principle, the full conditional distributions can be derived as follows:

$$p\left(\boldsymbol{\theta}_t \middle| \cdot\right) \propto \exp\left(-J_{\mathrm{e},t} - J_{\theta,t}\right) \tag{3.18}$$

$$\left(\boldsymbol{\mu}_\theta \middle| \cdot\right) \sim \mathcal{N}_0^{\boldsymbol{\mu}_\theta^{\mathrm{u}}}\left(\frac{1}{N_t}\sum_{t=1}^{N_t}\theta_t, \frac{1}{N_t}\Sigma_\theta\right) \tag{3.19}$$

$$\left(\boldsymbol{\mu}_{\mathrm{e}} \middle| \cdot\right) \sim \mathcal{N}\left(\frac{1}{N_t}\sum_{t=1}^{N_t}\mathrm{e}_t, \frac{1}{N_t}\Sigma_{\mathrm{e}}\right) \tag{3.20}$$

$$\left(\sigma^2_{\theta,p} \middle| \cdot\right) \sim InvGamma\left(\frac{N_t}{2} + \alpha_\theta, \frac{1}{2}\sum_{t=1}^{N_t}\left(\theta_{p,t} - \mu_{\theta_p}\right)^2 + \beta_\theta\right) \tag{3.21}$$

$$\left(\sigma^2_{\mathrm{e},p_{\mathrm{e}}} \middle| \cdot\right) \sim InvGamma\left(\frac{N_t}{2} + \alpha_{\mathrm{e}}, \frac{1}{2}\sum_{t=1}^{N_t}\left(e_{p_e,t} - \mu_{e_{p_{\mathrm{e}}}}\right)^2 + \beta_{\mathrm{e}}\right) \tag{3.22}$$

In Equation 3.19, $\mathcal{N}_a^b\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$ denotes a truncated Gaussian distribution with a mean $\boldsymbol{\mu}$ and a covariance matrix $\boldsymbol{\Sigma}$.

3.2.3. MAP Estimation of Parameters

Behmanesh et al. [15] proposed a simplified approach to estimating the maximum a posteriori (MAP) of parameters based on the conditional posterior distributions, so that the large computational costs from numerical simulation using Markov chain Monte Carlo (MCMC) sampling methods can be obviated. Basically, this approach iteratively estimates

the MAP of each parameter vector from the conditional posterior PDF (as shown in Equation 3.18 to Equation 3.22) until when the convergence criteria is reached. The detailed steps are:

(I) start with initial values of the parameter sets ${}^0\boldsymbol{\mu}_\theta$, ${}^0\boldsymbol{\mu}_{\mathrm{e}}$, ${}^0\boldsymbol{\Sigma}_\theta$, ${}^0\boldsymbol{\Sigma}_{\mathrm{e}}$, in which the left superscript ^{0} denotes initialization.

(II) at iteration j ($j = 1, 2, ...$), determine the MAP estimation of all the parameter sets from their conditional posterior distributions, that is:

$$(a) \quad {}^j\boldsymbol{\theta}_t = \underset{\boldsymbol{\theta}_t}{\operatorname{argmin}} \left({}^{j-1}\mathrm{J}_{\mathrm{e},t} + {}^{j-1}J_{\theta,t} \right)$$

$$(b) \quad {}^j\boldsymbol{\mu}_\theta = \frac{1}{N_t} \sum_{t=1}^{N_t} {}^{j-1}\boldsymbol{\theta}_t$$

$$(c) \quad {}^j\boldsymbol{\mu}_{\mathrm{e}} = \frac{1}{N_t} \sum_{t=1}^{N_t} {}^{j-1}\boldsymbol{e}_t$$

$$(d) \quad {}^j\sigma^2_{\theta,p} = \frac{\frac{1}{2}\sum_{t=1}^{N_t} \left({}^{j-1}\theta_{p,t} - {}^{j-1}\mu_{\theta_p} \right)^2 + \beta_\theta}{\frac{N_t}{2} + \alpha_\theta + 1}$$

$$(e) \quad {}^j\sigma^2_{\mathrm{e},p_{\mathrm{e}}} = \frac{\frac{1}{2}\sum_{t=1}^{N_t} \left({}^{j-1}e_{p_{\mathrm{e}},t} - {}^{j-1}\mu_{e_{p_{\mathrm{e}}}} \right)^2 + \beta_{\mathrm{e}}}{\frac{N_t}{2} + \alpha_{\mathrm{e}} + 1}$$

The MAP estimation in step (b) assumes that ${}^j\boldsymbol{\mu}_\theta$ lies within the boundaries prescribed in its prior (see Equation 3.9); otherwise, the closer boundary value should be used instead.

(III) check the convergence of the iteration in terms of parameter changes; if convergence is not reached, increase j by 1 and return to (II).

3.3. BMU-TL for Structural Identification and Damage Detection

This section establishes the framework of BMU-TL for structural identification and damage evaluation. Section 3.3.1 introduces the adaptation-regularization based transfer

learning (ARTL) used in this study to implement domain adaptation from the source domain containing numerical models to the target domain composed of experimental structures. Section 3.3.2 elaborates the framework of BMU-TL that incorporates TL guidance into the implementation of BMU. The established framework in this section will be used in Sections 3.4 and 3.5 for numerical and experimental case studies.

3.3.1. Adaptation-Regularization based Transfer Learning (ARTL)

Adaptation-regularization based transfer learning (ARTL) [76] is used in this study to demonstrate the effectiveness of incorporating domain adaptation into model updating. The general ARTL framework is supported by the structural risk minimization principle and the regularization theory. ARTL has three complementary functions: (1) minimizing the structural risk on the labeled data in source domain $\mathcal{D}_{\mathrm{s}}$; (2) minimizing the distribution difference between the joint probability distributions J_{s} and J_{t}; (3) minimizing the manifold discrepancy underlying the marginal distributions P_{s} and P_{t}.

Figure 3.2 illustrates the motivation of ARTL. Given a source domain $\mathcal{D}_{\mathrm{s}}$ with labeled data and a target domain $\mathcal{D}_{\mathrm{t}}$ with unlabeled data as shown in Figure 3.2(a), the classifier trained from $\mathcal{D}_{\mathrm{s}}$ cannot distinguish the data belonging to different classes in $\mathcal{D}_{\mathrm{t}}$ due to the substantial distribution difference. As shown in Figure 3.2(b), marginal distribution adaptation (MDA) reduces the distribution difference by drawing the moments of the two domains closer. Differing from MDA, conditional distribution adaptation (CDA) draws the intra-class centroids closer the inter-class centroids farther away and thus largely improves the performance of the trained classifier as shown in Figure 3.2(c). Figure 3.2(d) shows that manifold regularization (MR) realigns the hyperlane to maximize the manifold consistency underlying the marginal distributions.

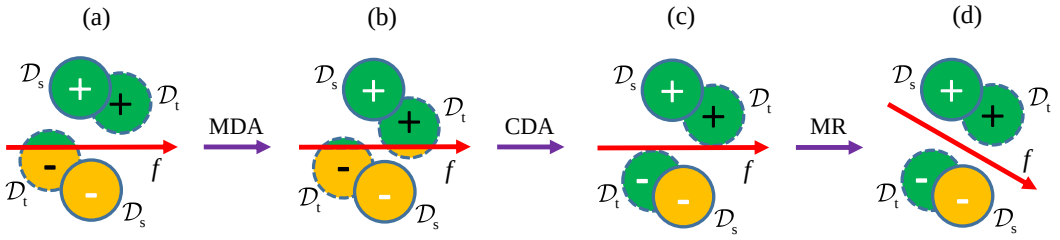


Figure 3.2. Motivation of ARTL (reproduced from [76]). f is the hyperlane of the trained classifier; $\mathcal{D}_\mathrm{s}$ is the source domain; $\mathcal{D}_\mathrm{t}$ is the target domain; "+" and "-" denote two different class labels; MDA denotes marginal distribution adaptation; CDA denotes conditional distribution adaptation; MR denotes manifold regularization. (a) the original domains and the trained classifier before ARTL; (b) effects after MDA; (c) effects after MDA and CDA; (d) effects after MDA, CDA, and MR.

Suppose the classifier has a prediction function $f = \boldsymbol{w}^\mathrm{T}\phi(x)$, where $\boldsymbol{w}$ contains the classifier parameters, x is the feature data, and $\phi : \mathcal{X} \mapsto \mathcal{H}$ is the feature mapping function that projects the feature vector from its original space to the Hilbert space $\mathcal{H}$. Then the learning framework of ARTL is formulated as

$$f = \arg\min_{f \in \mathcal{H}_K} \sum_{i=1}^{n} \ell\left(f(\mathbf{x}_i), y_i\right) + \sigma\|f\|_K^2 + \lambda D_{f,K}\left(J_\mathrm{s}, J_\mathrm{t}\right) + \gamma M_{f,K}\left(P_\mathrm{s}, P_\mathrm{t}\right) \tag{3.23}$$

The sum of first two terms on the right hand side of Equation 3.23 is the structural risk functional defined following the structural risk minimization principle, that is

$$f = \arg\min_{f \in \mathcal{H}_K} \sum_{i=1}^{n} \ell\left(f(\mathbf{x}_i), y_i\right) + \sigma\|f\|_K^2 \tag{3.24}$$

in which $\ell()$ is the loss function measuring the fitness of f when used for estimating the labels of $\mathcal{D}_\mathrm{s}$, $\|f\|_K^2$ is the squared norm of f in $\mathcal{H}_K$, with K being the kernel function induced by ϕ such that $\langle\phi(\mathbf{x}_i), \phi(\mathbf{x}_j)\rangle = K(\mathbf{x}_i, \mathbf{x}_j)$. However, the classifier inferred by Equation 3.24 cannot generalize well to the target domain $\mathcal{D}_\mathrm{t}$, as it requires that the training and testing data originate from the same probability distribution. Thus, ARTL entails minimizing the distribution distance between the joint distributions J_s and J_t, that is the

$D_{f,K}(J_s, J_t)$ regularized by λ. On the other hand, considering that the unlabeled data may have the potential to reveal the underlying truth in the target domain, the marginal distribution P_s and P_t can be used for minimizing the discrepancy between the predictive structure of f and the intrinsic manifold structure of data, by adding the manifold regularization term $M_{f,K}(P_s, P_t)$ with the regularization parameter γ. More details about the ARTL can be found in Reference [76].

3.3.2. BMU-TL

Introducing pattern recognition to guide the process of Bayesian model updating, this study proposes BMU-TL to solve the modeling uncertainty problem in model-based SHM. Figure 3.3 illustrates the framework of BMU-TL for structural identification and damage detection. Given a certain structure for monitoring, an initial numerical model M_{00} can be established from the design parameter and some simplification/assumption based on our understanding of the structural characteristics. M_{00} corresponds to the intact state of the structure, that is S_0. Significant modeling error from improper simplification or assumption may be introduced into this modeling process, which causes large deviation of the model's behavior and thus the estimated features from the target structure. Then BMU using the measured features yields a model M_0 with updated stiffness parameter $\boldsymbol{\theta}_0 = \{\theta_1, \theta_2, ...\theta_N\}$. $\boldsymbol{\theta}_0$, though updated, may not correspond to the true values of structural stiffness of S_0, as model updating cannot intrinsically correct the modeling error. When damage happens to the structure at certain location(s), it becomes a different structure S_1. In traditional model updating methods, the numerical model M_0 is updated via BMU using the measured features from S_1, yielding an updated numerical model denoted as M_1' having parameter $\boldsymbol{\theta}_1'$; then damage locations and severities can be identified

through comparing $\boldsymbol{\theta}_1'$ with $\boldsymbol{\theta}_0$. However, the identified damage may not reflect the real structural health condition due to modeling errors contained in establishing and updating the numerical model. In the proposed BMU-TL method, instead of implementing BMU directly, the potential damage location can be identified first through pattern recognition via ARTL taking the numerical models as lying in the source domain and experimental structures as in the target domain. With the recommended damage sites from ARTL, that is $L = [l_1, l_2, ...l_n]$, the numerical model is updated from M_0 with only the parameters corresponding to L updated, yielding an updating model M_1 with parameter $\boldsymbol{\theta}_1$. Then the damage severities at site L, $D = [d_1, d_2, ...d_n]$, can be obtained through comparing $\boldsymbol{\theta}_1$ with $\boldsymbol{\theta}_0$.

It should be noted that domain adaptation does not intrinsically correct modeling error in model updating. Hence, the identified damage severities through BMU-TL may still be inaccurate especially in the presence of severe modeling error. However, BMU-TL can accurately detect the correct damage locations, which saves a large amount of work in practical application for retrofitting and maintenance. Additionally, ARTL reduces the dimension of target parameter to be updated via BMU and thus has the potential of saving the computational costs of model updating. Moreover, through combining pattern recognition with model updating, BMU-TL has the potential of directing the process of structural identification and damage detection, so that the SHM outcome does not violate fundamental physical laws.

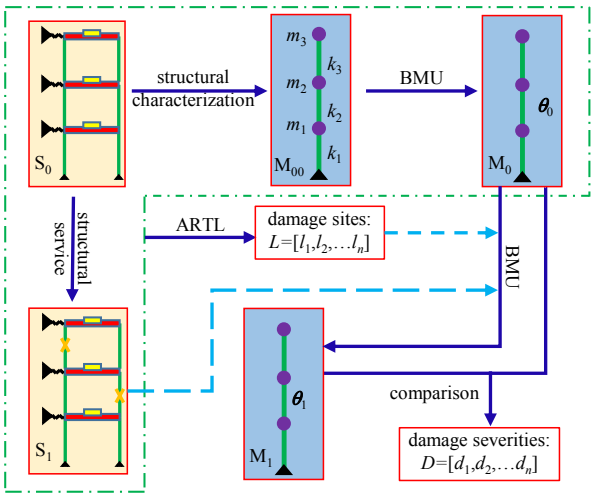


Figure 3.3. Framework of transfer learning aided Bayesian model updating (BMU-TL) and damage identification. S_0 denotes the experimental structure before damage happens; S_1 denotes the structure when damage happens to the columns labeled with crosses; M_{00} is the initial numerical model established from engineering drawings before model updating; M_0 is the model with the stiffness parameter $\boldsymbol{\theta}_0$ updated using the measured features from S_0 via hierarchical Bayesian model updating (BMU); M_1 is the model with parameter $\boldsymbol{\theta}_1$ updated using measured features from S_1 via BMU-TL.

3.4. Numerical Case Study

This section implement a numerical case study to validate the effectiveness of BMU-TL in structural identification and damage evaluation when compared with BMU. Section 3.4.1 introduces the structural model used in this numerical study, followed by the presentation of results and discussions in Section 3.4.2.

3.4.1. Structural Model for Numerical Analysis

This numerical analysis uses an 8-floor structure as shown in Figure 3.4 with the corresponding parameters listed in Table 3.1. Modeling error is introduced to the numerical model by omitting the masses of nonstructural components and lateral supports, as shown in Figures 3.4 (a) to (c). This analysis assumes that each floor stiffness follows a Gaussian distribution with the mean value and standard deviation (std) listed in Table

Table 3.1. Structural parameters of the numerical model and its natural frequencies before damage happens.

item	value, $\times m_0$	item	value, $\times m_0$	item	mean value, $\times k_0$	std, $\times k_0$	item	value, $\times k_0$	item	value, Hz
m_1	1.0	dm_1	0	k_1	1.0	0.05	k_{b1}	0	f_1	2.82
m_2	1.0	dm_2	0.3	k_2	2.0	0.10	k_{b2}	0	f_2	7.90
m_3	0.5	dm_3	0	k_3	3.0	0.10	k_{b3}	0	f_3	14.26
m_4	0.5	dm_4	0.1	k_4	4.0	0.10	k_{b4}	1.0	f_4	19.57
m_5	0.5	dm_5	0.1	k_5	2.5	0.10	k_{b5}	0	f_5	26.22
m_6	0.5	dm_6	0	k_6	3.5	0.10	k_{b6}	0.8	f_6	35.20
m_7	1.0	dm_7	0.35	k_7	1.5	0.05	k_{b7}	0.3	f_7	37.95
m_8	0.5	dm_8	0	k_8	5.0	0.10	k_{b8}	0	f_8	43.61

m_0: reference floor mass, 2.35 kg; k_0: reference column stiffness, 7479.11 N/m.

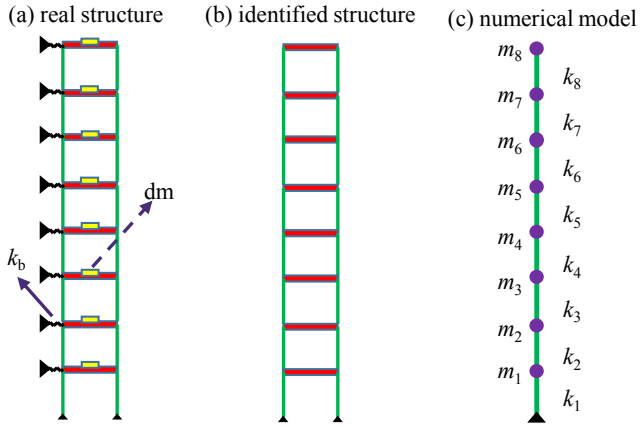


Figure 3.4. The 8-floor structure and numerical model in Bayesian model updating with modeling error. (a) the real structure of SHM interest; dm denotes the mass of unknown nonstructural components on each floor; k_b denotes lateral supports from adjacent buildings or temporary braces. (b) the identified structural form from engineering drawings, which omits dm and k_b and thus contains modeling error. (c) the established numerical model from the identified structure in (b); m_i denotes the floor masses and k_i the column stiffness with $i = 1, 2, ..., 8$.

3.1 to simulate the uncertainty caused by ambient/environmental factors. The natural frequencies of the intact structure as listed in Table 3.1 are calculated using the mean values of the stiffness parameters. Damage can be introduced via reducing the mean value of certain column stiffness parameters by a certain portion, which will cause variations of the measured modal properties. 5% measurement noise is added to the numerically calculated

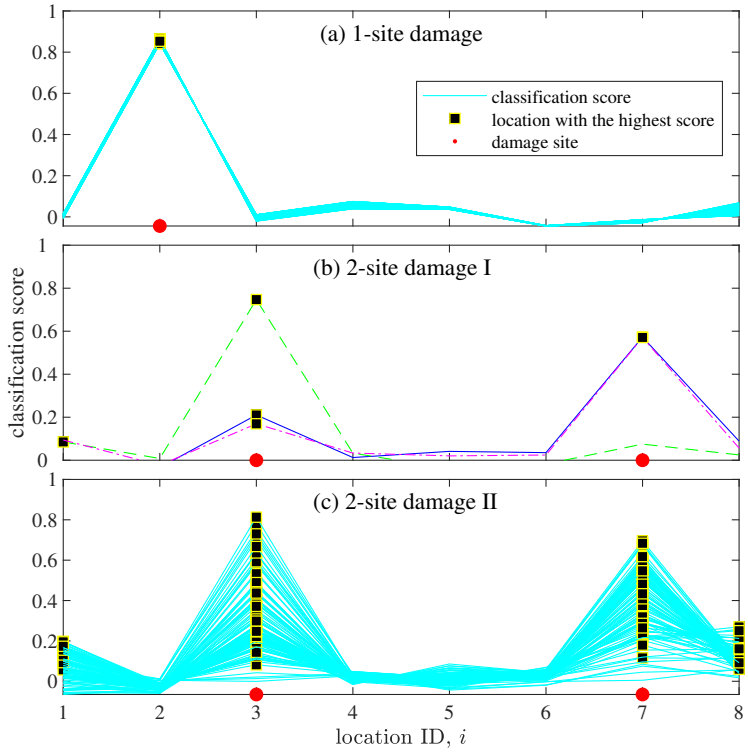


Figure 3.5. Classification scores in two damage cases from ARTL. (a) 100 example 1-site damage cases at k_2 with various damage severities. (b) 3 example 2-site damage cases at k_3 and k_7. (c) 100 example 2-site damage cases at k_3 and k_7.

modal properties to resemble the effects of measurement uncertainty. This case study implements 50 simulated tests (i.e., $N_t = 50$ as defined in Section 3.2.2) to generate the measured data for model updating that include the eigenfrequencies and mode shapes of the first four modes. Following the author's previous research [8, 77], the normalized frequency change ratio (NFCR) and mode shapes ($\mathbf{\Phi}$) are used as features for pattern recognition in TL, that is $x = [\text{NFCR } \mathbf{\Phi}]$ where x is the feature vector defined in Section 3.3.1.

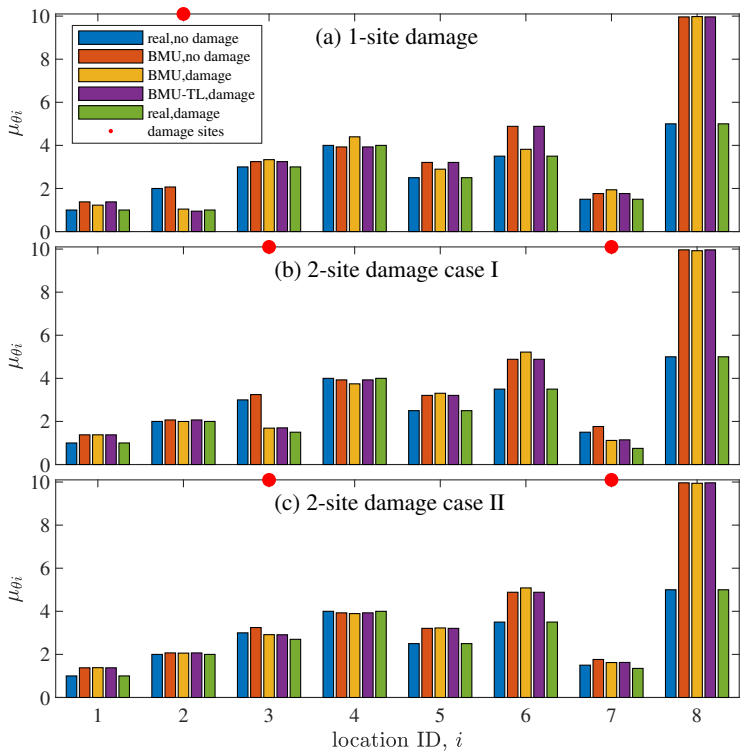


Figure 3.6. Identified structural stiffness coefficients using BMU and BMU-TL and the target values. (a) 50% damage on k_2. (b) 50% damage on k_3 and k_7. (c) 10% damage on k_3 and k_7.

3.4.2. Results and Discussions

This section presents the results of BMU-TL for structural identification and damage detection using the example structure introduced in Section 3.4.1 and compares with that of BMU without TL regarding the accuracy of damage localization and quantification. This section first considers one single site damage case with 50% damage on k_2 and two 2-site damage on k_3 and k_7 with different damage severities (50% and 10% respectively), and then presents the results of all possible 1-site damage cases with 50% damage severity, leaving multi-site damage cases for future work considering the challenge of pattern recognition for multi-site damage detection.

To demonstrate the advantage of applying TL for damage localization, Figure 3.5

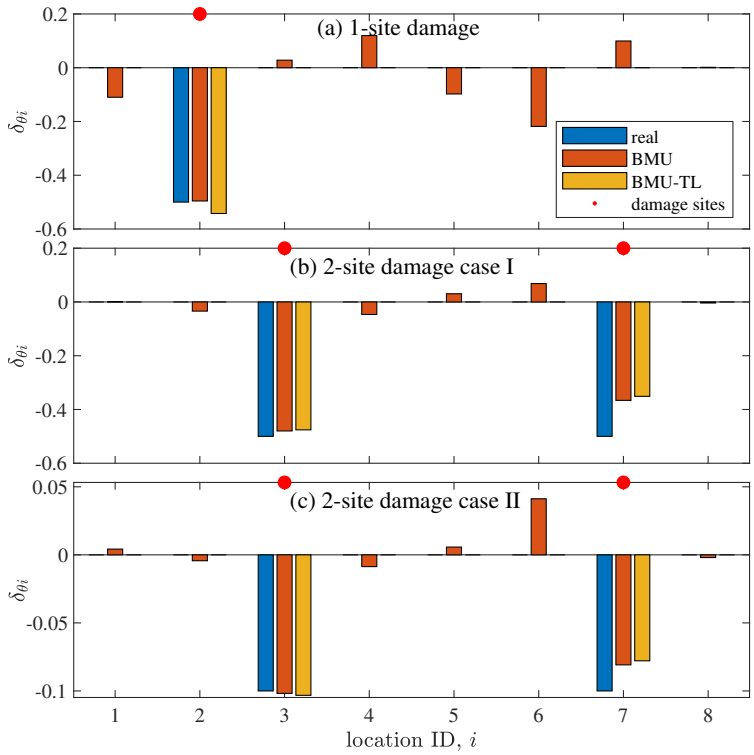


Figure 3.7. Identified structural damage severities using BMU and BMU-TL and the target values. (a) 50% damage on k_2. (b) 50% damage on k_3 and k_7. (c) 10% damage on k_3 and k_7.

first shows the classification scores produced by ARTL for two damage cases with randomly varied damage severities from 10% to 90%. In damage detection using machine learning classification, each class corresponds to a certain damage location, and the classifier recommends the location with the highest score. Figure 3.5 (a) shows the output scores for 100 1-site damage cases with various damage severities on k_2. It shows that ARTL always recommends the correct damage location with the classification score for location 2 much higher than that elsewhere. In comparison, if without TL, the trained classifier using the data from the source structure only correctly predicts less than 70% damage locations. This approves the effectiveness of TL for recommending potential damage location when the established model contains significant modeling uncertainty. Figure

3.5 (b) shows the classification scores for three examples cases of 2-site damage cases with damage on k_3 and k_7, and Figure 3.5 (c) shows all the 100 example cases in this damage scenario. It can be seen that ARTL yields the highest two scores on the target damage locations, which happens to 74% of all cases in this damage scenario. This finding indicates that ARTL has the potential of helping identify the correct damage locations when multi-site damage happens, though accurate damage localization remains challenging which may entail TL based on multi-label classification.

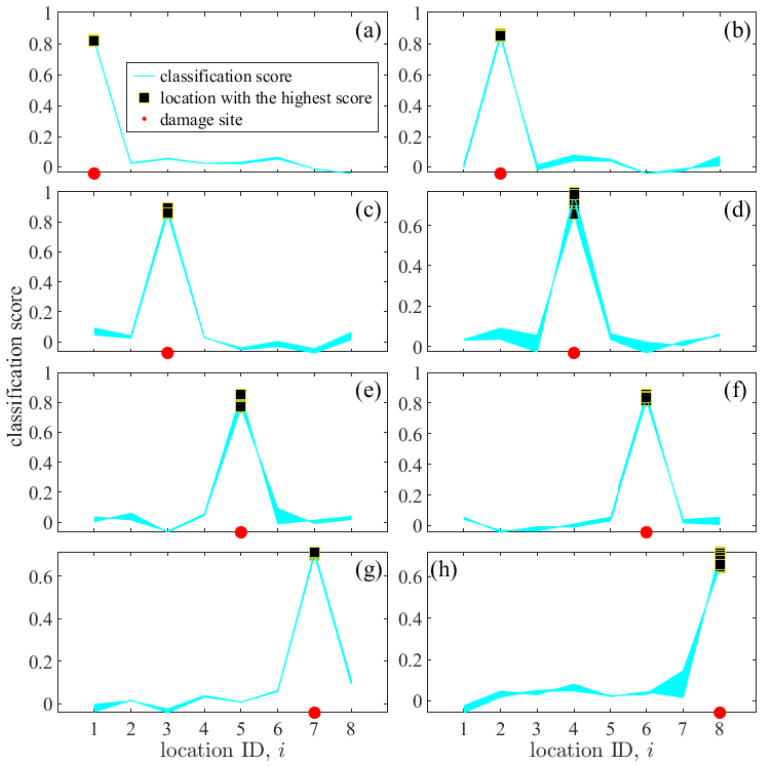


Figure 3.8. Classification scores for 1-site damage cases from ARTL. (a) to (h) correspond to damage case I to VIII with 50% damage at k_1 to k_8, respectively.

Figure 3.6 compares the updated parameter mean values (μ_{θ_i} $(i = 1, 2, ..., 8)$) of the intact structure and damaged structure from BMU and BMU-TL with the real values. The three damage cases mentioned above are considered here assuming that TL helps

identify the correct damage locations. It shows that because of modeling error, the updated structural parameters significantly deviate from the real values. This deviation is especially obvious at location 8, most probably because varying θ_8 has the largest effects in compensating the influence of modeling error on measured features. When damage happens, BMU without ARTL can recommend incorrect damage sites as well as severities. Moreover, the updated structural parameter values can become larger than that from the updated intact structure, which indicates structural strengthening and thus is misleading for structural retrofitting and maintenance. In comparison, TL recommends the correct damage location and thus avoids incorrectly updating the structural parameters corresponding to no damage.

Figure 3.7 compares the identified damage severities (δ_{θ_i} $(i = 1, 2, ..., 8)$) from BMU and BMU-TL with the real values. Without constraining the possible damage locations, BMU updates structural parameters in a way minimizing the discrepancy of modal features and thus compensating the effects of modeling error. As a result, the updated parameters values may not reflect the true structural condition especially regarding damage locations. For example, as shown in Figure 3.7 (c), when 10% damage happens to k_3 and k_7, BMU without TL recommends 4% damage at k_6 and slight damage elsewhere other than the target sites.

Considering the challenge of identifying multi-site damage locations and corresponding damage severities, the rest of this section focuses on 1-site damage cases and further compares the performance of BMU and BMU-TL in structural damage detection. Figure 3.8 presents the classifications scores from ARTL for the eight damage cases (I, II, ... VIII) that have 50% damage on k_1, k_1,..., k_8, respectively. It shows that ARTL recom-

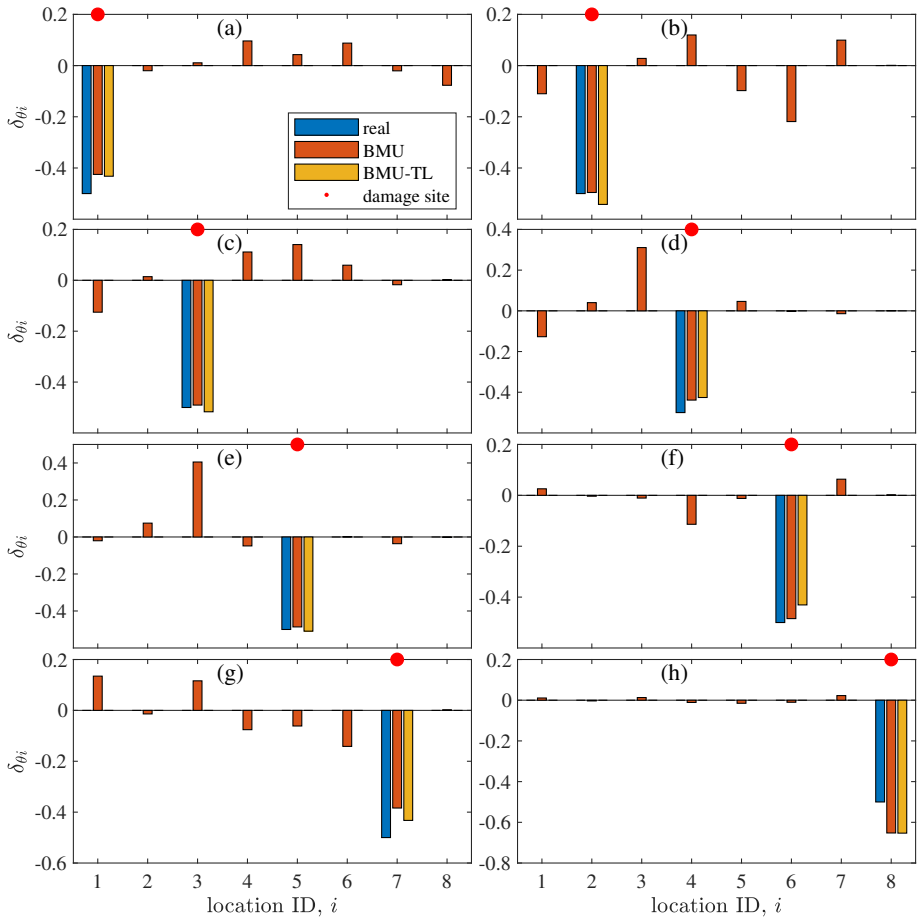


Figure 3.9. Identified structural damage severities for 1-site damage cases using BMU and BMU-TL and the target values. (a) to (h) correspond to damage at k_1 to k_8 respectively.

mends the correct damage location for all cases with a classification score at the target location much higher than that elsewhere. This encouraging outcome from ARTL guarantees correct damage localization when 1-site damage happens.

Figure 3.9 compares the identified damage severities with the real values for damage case I to VIII. Again, BMU without TL yields misleading results of damage identification when compared with BMU-TL. Interestingly, when damage happens at k_8 in case VIII, BMU identifies negligible incorrect damage elsewhere. As shown in Figure 3.6, modeling error causes the largest deviation in estimating the parameters for k_8. Hence, when

damage indeed happens at k_8, model updating allows compensating the effects of modeling error largely by adjusting the value of parameter for k_8. Figure 3.10 compares the sum of mis-identified damage severities at all locations (E_δ) using BMU and BMU-TL for cases I to VIII. It shows that BMU-TL has much smaller identification error sum than BMU for all the eight cases. E_δ from BMU-TL is below 0.20 for all the cases and below 0.10 for most cases. In comparison, E_δ from BMU is always higher than 0.20 and sometimes larger than 0.50. This comparison further highlights the improvements that TL brings to damage identification in model updating.

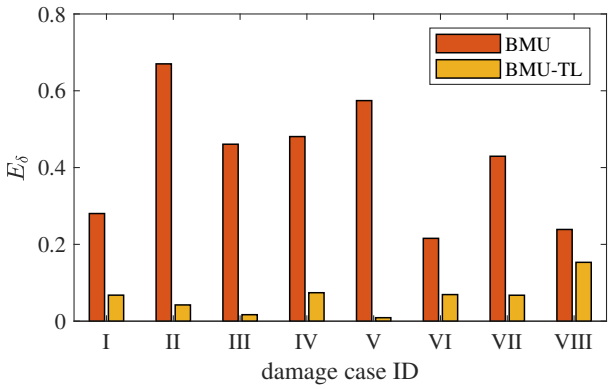


Figure 3.10. Sum of incorrectly identified damage severities using BMU and BMU-TL. Damage cases I to VIII has 50% damage at k_1 to k_8, respectively.

3.5. Experimental Validation

An experimental study is conducted in this section to further validate the effectiveness of BMU-TL in structural damage identification. Figure 3.11 shows a 6-floor building model used in this study. The floors are made of steel plates whose dimensions are: $11.875 \times 3 \times 0.5$ in.^{3} for floors 1 and 2; $11.875 \times 3 \times 0.25$ in.^{3} for floors 3 to 6. Each floor contains two aluminum columns with identical dimensions, that is $14 \times 3 \times 0.125$ in.^{3}.

As shown in Figure 3.11, columns are fixed to floor plates through bolts at the plate ends, and the whole model is fixed to a shaking table. Accelerometers are placed on each floor to measure the structural responses under ground motions. To introduce damage to the intact structure in a nondestructive manner, variable masses are added to a certain floor to indirectly simulate structural damage. With the intact case denoted as D0, symbols D1 to D6 denote cases with additional masses added to floors 1 to 6. Each damage case contains 10 subcases in which the ratio of the added mass and the original mass ranges from 0.1 to 1.0.

White noise excitations are used to simulate service loading applied to a real building. Ten random excitations are generated and used in the intact case and each damaged subcase. Hence, a total number of $(6 \times 10 + 1) \times 10 = 610$ tests are run for 7 damage classes (i.e., D0 to D6), which provides sufficient data for pattern recognition in TL and comprehensively examining the effectiveness of TL guidance in improving the BMU performance. Figure 3.12 shows the measured responses of the intact model under certain white excitation with a sampling rate of 2 kHz.

Modal properties including frequencies and mode shapes are extracted from the measured responses using the frequency domain decomposition (FDD) [36] method of operational modal analysis (OMA). Before implementing OMA, the measured responses are processed with a Hanning window to reduce information leakage and a digital filter to exclude the signal components outside the model's frequency range. Figures 3.13 and 3.14 compare the extracted frequencies and mode shapes of the intact model with that calculated through eigenanalylsis with respect to the model mass and stiffness matrices. Figures

Figure 3.11. A 6-floor frame model used for experimental study.

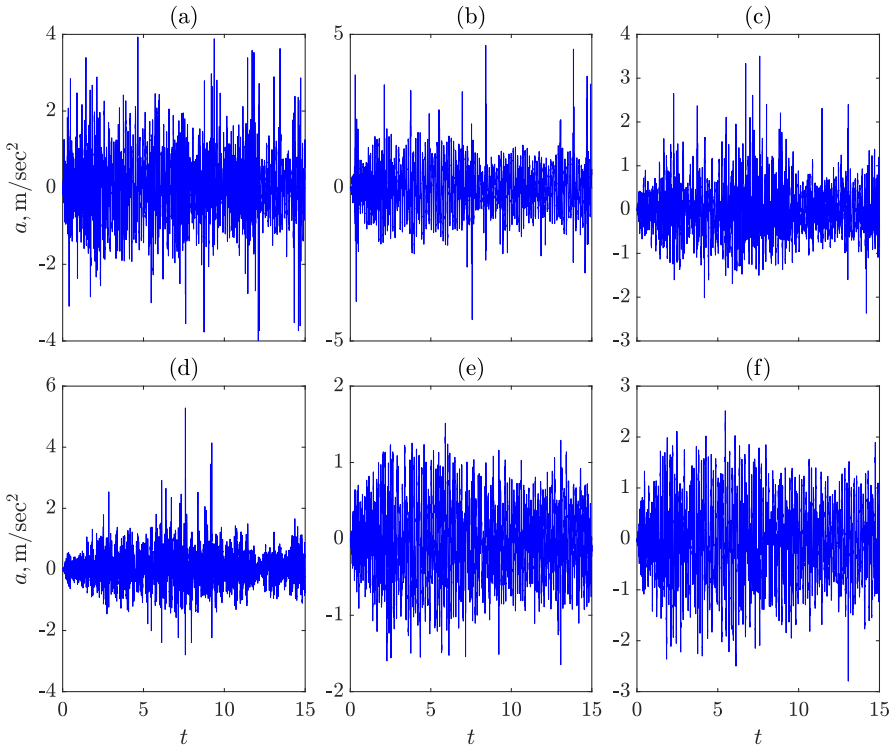


Figure 3.12. Measured accelerations of each floor under white noise excitation of the intact condition. (a) to (f): responses of floors 1 to 6.

3.13 and 3.14 show that the modal properties from OMA are consistent with the theoretical results, which confirms the effectiveness of the OMA technique and validates the established theoretical model. The observed difference can be attributed to the possible limitation of OMA, simplification and approximation of the physical structure in establishing the numerical model, and disturbance from measurement and system noise. Additionally, Figure 3.14 (f) shows that the OMA performance might be limited for identifying higher modes information. In this section, as done in the numerical study, properties of the first three modes are selected for model updating and pattern recognition.

Figures 3.15 and 3.16 compare frequencies and mode shapes of all damage cases,

respectively. It can be seen from Figures 3.15 (a) to (f) that frequencies do not vary consistently when damage happens to a certain floor. For example, when damage severity increases, frequencies f_2, f_3, and f_4 of D1 cases reduce monotonically while other frequencies do not vary in this manner. Additionally, many outliers can be observed in the plot of f_6, showing the limited OMA performance when extracting higher-order modal properties. Similar to the variation of frequencies due to various damage cases, mode shapes show different variations when certain damage happens, as can be observed from the deviation of damaged mode shapes from the intact case, as shown in Figure 3.16. In Figure 3.16, when damage happens to the 4^{th} floor, $\mathbf{\Phi}_1$ in (d1) has larger variation than $\mathbf{\Phi}_2$ in (d2) and $\mathbf{\Phi}_3$ in (d3), especially at the 1^{st} floor. Figures 3.15 and 3.16 demonstrate that neither a certain mode nor multiple modes together can be used to directly localize or quantify structural damage. Hence, feature extraction and pattern recognition are necessary for achieving desired performance of damage identification.

As done in the numerical study, NFCR and $\mathbf{\Phi}$ are used as features for pattern recognition in this section. To visualize the effectiveness of extracted features in distinguishing examples belonging to different damage classes, Figure 3.17 scatters the instances belonging to damage cases D0 to D6 using three principal components (PCs) of features obtained from principal component analysis (PCA). It can be observed that some instances belonging to different damage cases can be well separated in certain feature space while they may overlap in other spaces. For example, some D3 cases overlap with D1 cases in the PC1, PC2 space as shown in (a); however, they can be well distinguished from each other when plotted in other two dimensional PC spaces as shown in (b) and (c). Therefore, it can be concluded that examples belonging to different cases should be more

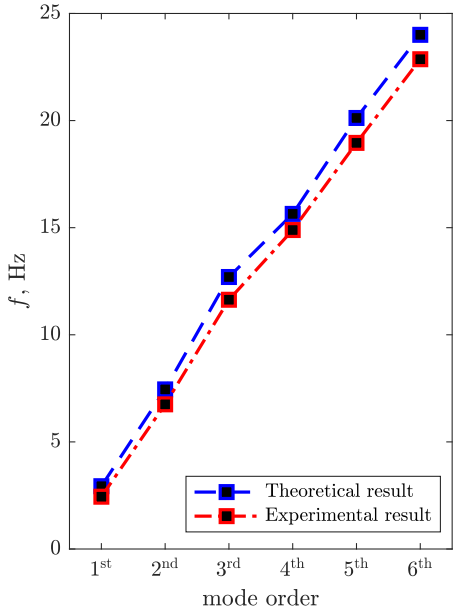


Figure 3.13. Frequencies obtained from the experiment and theoretical model

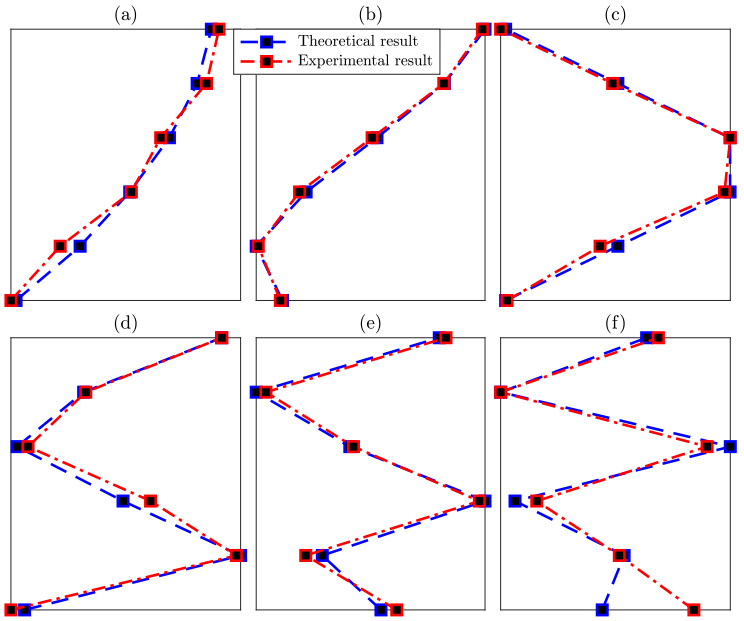


Figure 3.14. Mode shapes from the experiment and theoretical model. (a) to (f) show the shapes of modes 1 to 6 (i.e., $\mathbf{\Phi_1}$ to $\mathbf{\Phi_6}$), respectively.

distinguishable in the non-reduced full feature space.

The extracted damage sensitive features are then input into the TL framework for damage identification, which yields an accuracy of 95.2% regarding damage localization on the experimental data. As shown in Figure 3.18, all identification errors happen in slight damage cases. This phenomenon happens because slight damage usually does not cause as considerable variation of modal properties as moderate or severe damage does (Figure 3.15). Despite this challenge, most slight damage cases are categorized into the correct damage class through TL, which is encouraging in data-driven damage identification with most measured cases unlabeled.

Before proceeding to improve the BMU performance with TL, this study first updates floor stiffness parameters with correct and incorrect floor masses used in the reference model to have an overview of the effects of modeling error in model updating. Figure

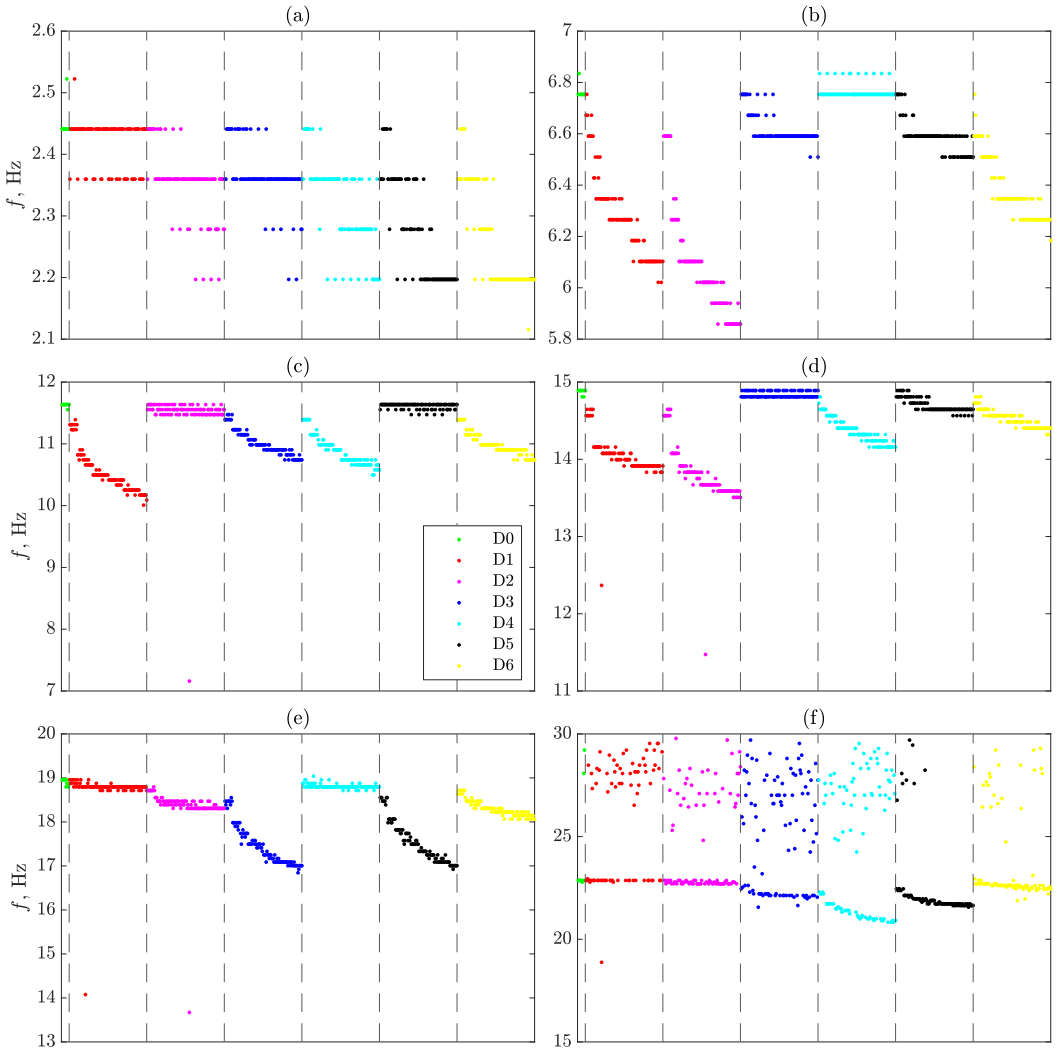


Figure 3.15. Variations of frequencies when damage happens. (a) to (f) show the 1st to 6th mode frequency, namely, f_1 to f_6. D0 stands for the intact case with no mass addition on any floor. Di represents cases with mass added on the i^{th} ($i = 1, 2, ..., 6$) floor. Within each damaged case, the damage severity increases from the left to the right of plot.

3.19 shows the results of BMU on floor stiffness coefficients with structural masses set as (1) the measured quantities of the intact model (2) the measured quantity of the corresponding cases. Scenario (1) contains explicit modeling error from incorrectly measured floor masses for damaged cases D1 to D6, and scenario (2) approximately corrects this modeling error with the measured floor masses. It shows that the updated stiffness parameters of damaged cases (i.e., D1 to D6) are considerably deviant from the real values

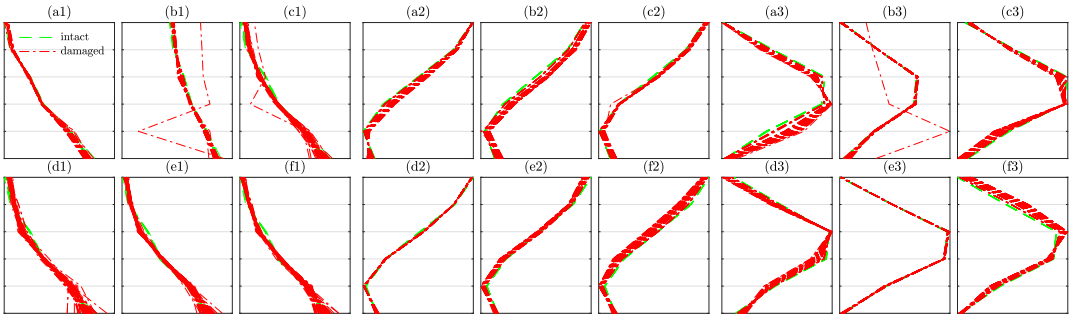


Figure 3.16. Variations of mode shapes when damage happens compare with intact cases. (a1) to (f1) show the 1st mode ($\mathbf{\Phi}_1$) when damage happens on floors 1 to 6, respectively; (a2) to (f2) show the 2nd mode ($\mathbf{\Phi}_2$); (a3) to (f3) show the 3rd mode ($\mathbf{\Phi}_3$).

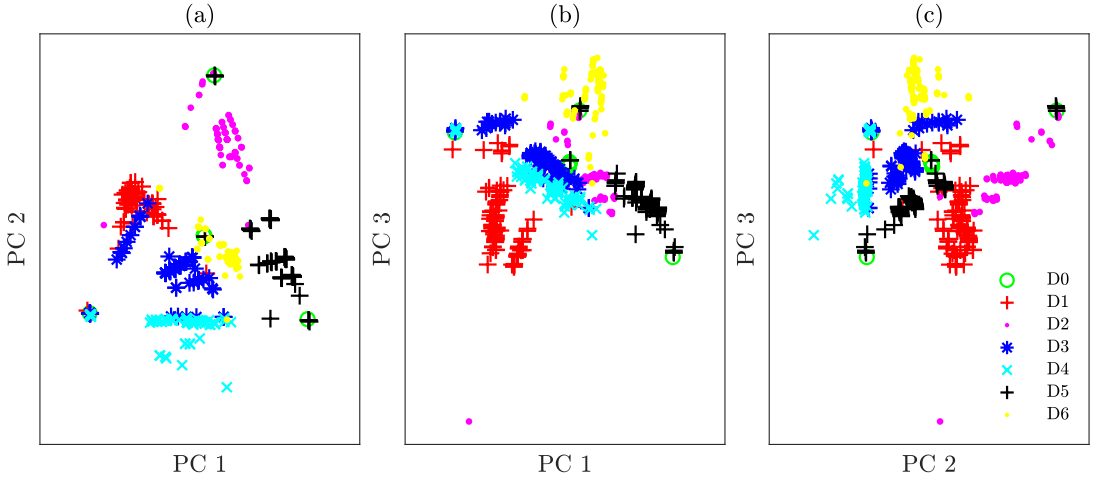


Figure 3.17. Distribution of instances in 2D PCA-reduced feature space. Three combinations of PCs are used to visualize the effectiveness of extracted features in distinguishing examples of different damage cases.

(i.e., that of D0 cases) as a result of modeling error. With modeling errors approximately corrected by using measured floor masses of each case, the updated parameter values are drawn much closer to that of D0. As the added masses of damaged cases are hard to be measured accurately (i.e., still slight modeling error of floor mass exists in each damaged case), the corresponding BMU results are still deviant from that of D0.

Figure 3.20 shows the identified damage severities from BMU and BMU-TL together with the target values. It shows that with incorrect damage locations excluded by

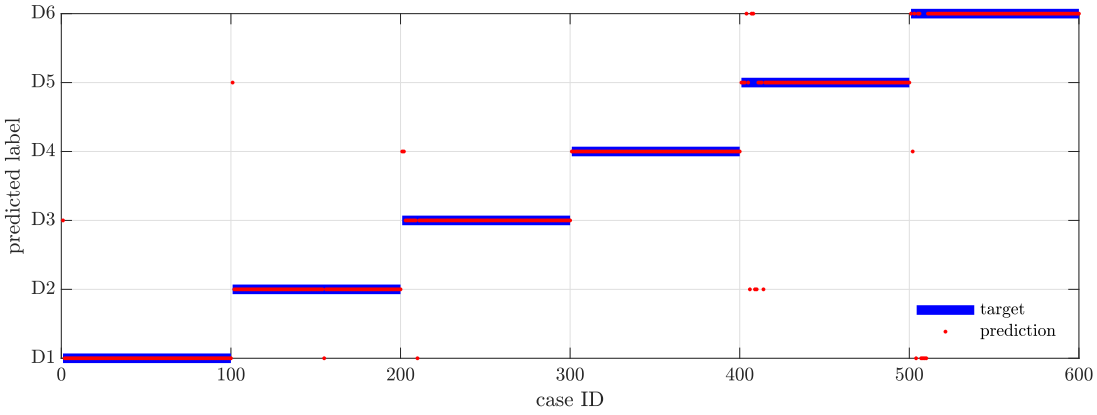


Figure 3.18. Results of transfer learning. Each damage class contains 100 cases, that is D1 corresponds to cases 1-100, D2 corresponds to 101-200, etc. Within each class, the damage severity get larger as the case ID increases.

TL guidance, BMU yields generally more accurate damage quantities for most damage cases. Considering the possible error when calculating the target damage severities as denoted by the green lines, the slight difference between BMU-TL results and the target values are acceptable. Hence, this section further validates the effectiveness of TL in guiding BMU and improving the performance of damage identification.

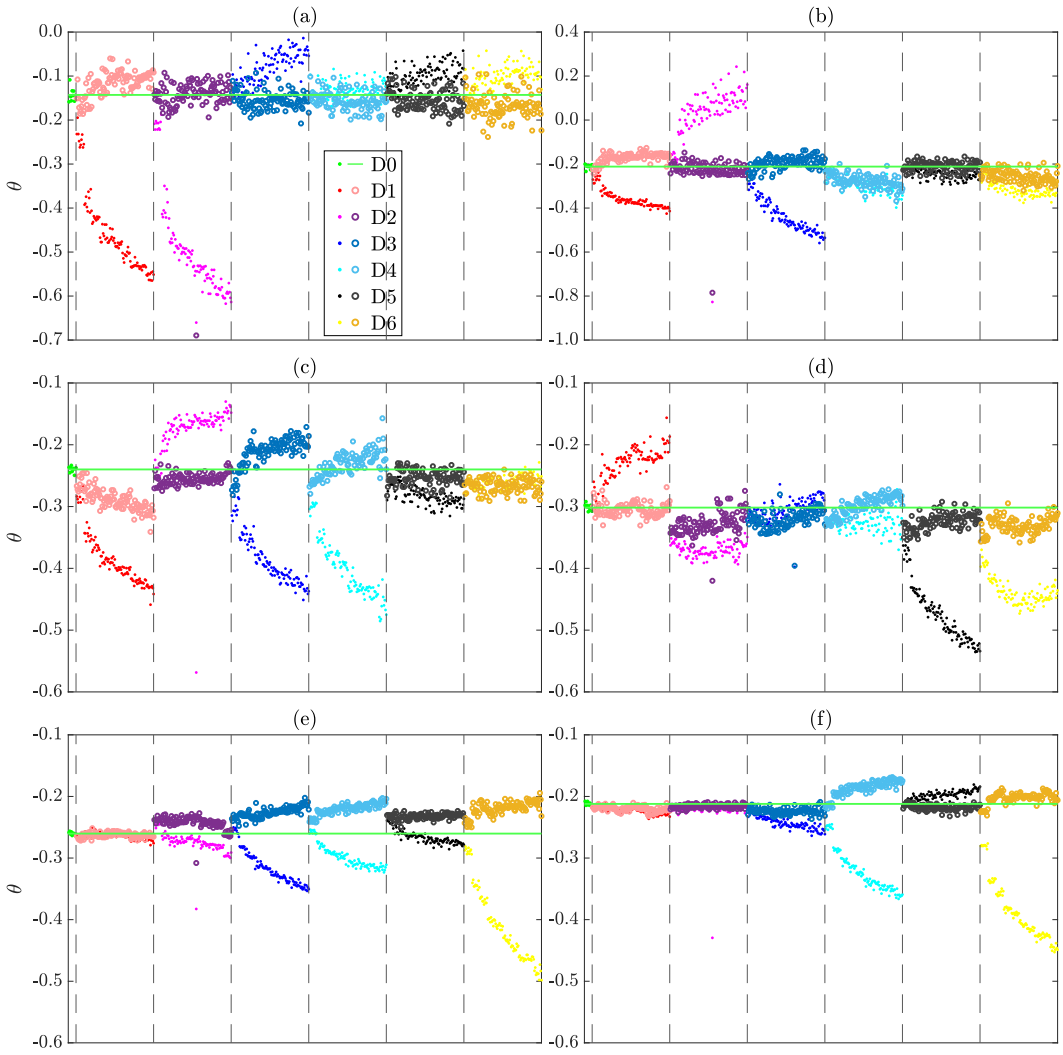


Figure 3.19. BMU results on structural stiffness coefficients, indicating the effects of modeling error. Scatter plots with dot markers (.) stand for results of BMU with modeling error from incorrect floor masses, and those with circle markers (o) represent results with corrected floor masses. The intact cases (D0) contain no mass error, and the green horizontal line represents the average BMU result of intact cases.

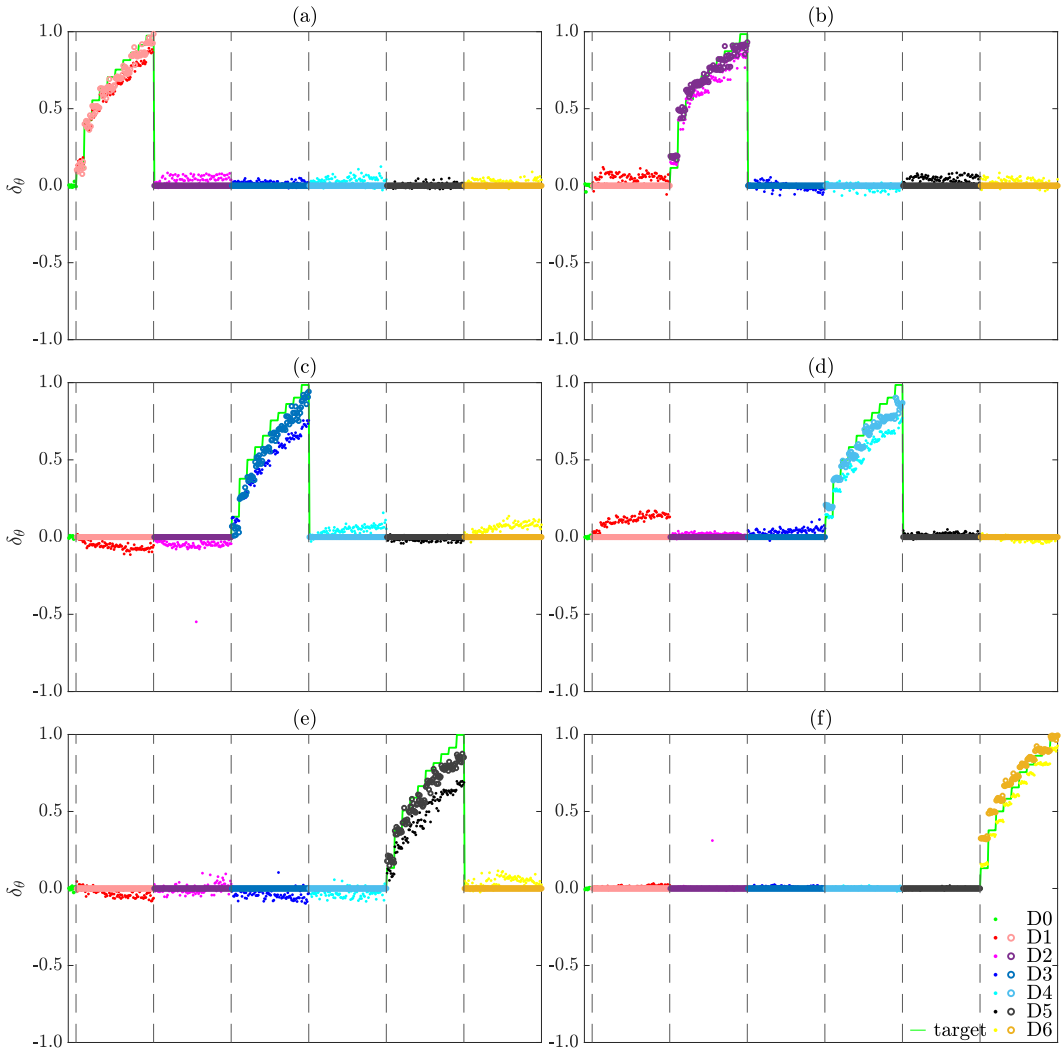


Figure 3.20. Identified structural damage severities using BMU and BMU-TL in comparison with the targets. Scatter plots with dot markers (.) stand for results of BMU without TL guidance, and those with circle markers (o) represent results of BMU-TL.

3.6. Summary

This study attempts to improve the performance of Bayesian finite element model updating when significant modeling uncertainty/error exists in establishing the numerical models. A domain adaptation based transfer learning technique is used to guide the structural parameter optimization in model updating and circumvent the influence of modeling error in damage localization. To evaluate the performance of the proposed method, this study conducts a numerical case study using an 8-DOF system and an experimental case study using a 6-floor building structure. Domain adaptation is realized with the adaptation-regularization based transfer learning method that minimizes the influence of domain difference in pattern recognition through regularizing the distribution and manifold discrepancy. The results from both numerical and experimental case studies show that transfer learning can help identify the correct damage locations and thus direct and regularize the model updating process. As a result, the proposed method can identify damage locations and severities more accurately. In detail, transfer learning in the numerical case study yields an accuracy of 100% of damage locations for single-site damage cases. With guidance from transfer learning, Bayesian model updating reduces the sum of errors of updated structural parameters largely from more than 50% to lower than 20%. In the experimental study, the structural damage is introduced through adding pieces of variable masses to a certain floor. Bayesian model updating on structural stiffness coefficients demonstrates the effects of modeling uncertainty on structural identification and damage evaluation. Subsequently, with the extracted damage sensitive features, transfer learning identifies damage locations with an accuracy of 95.5%. Furthermore, with trans-

fer learning guidance, Bayesian model updating identifies damage severities much closer to the real values than without transfer learning guidance. To sum up, this study provides a new method for incorporating pattern recognition into model updating, or combining data-driven SHM with model-based SHM. The proposed new method has the potential of benefiting the SHM community with the merits of data-driven and model-based methods.

Chapter 4. Multi-Site Structural Damage Identification Using a Multi-Label Classification Scheme of Machine Learning

4.1. Introduction

Chapters 4 and 5 will deal with the multi-site damage identification problem in data-driven SHM. Accurate identification of multi-site damage cases requires resolving the inter-correlations between damage cases, especially those sharing common damage sites. To this end, the multi-label classification method of machine learning is investigated in the present chapter.

Multi-site structural damage identification is significantly important in maintaining the safety and reliability of engineering structures. However, it has received insufficient research efforts so far due to its complexity and difficulty [5]. In comparison, multi-site damage identification is challenging as the multi-site damage occurrence can enlarge the adverse influence of uncertainties (measurement, model and etc.) on damage indicators or derived features. Consequently, effective methods for single-site damage identification will probably lose their effectiveness in identifying multiple damage sites.

Approaches to structural health monitoring generally fall into two categories, i.e., model-based approaches and data-based approaches. Model-based approaches identify structural damage using a physics-based model usually with recourse to finite element (FE) model-updating, through minimizing the residuals between measurements and simulated results [79, 80]. Model-based methods necessitate an accurate numerical model and repetitive model-updating, both of which are computationally demanding. In addition,

Reprinted from Measurement, Volume 154. Zhiming Zhang and Chao Sun, Multi-Site Structural Damage Identification Using a Multi-Label Classification Scheme of Machine Learning, 107473, Copyright (2020), with permission from Elsevier.

model-based methods are usually subject to the uncertainty arising from the physical variability, experimental validity, and model-form error [5]. Though with inevitable shortcomings, model-based methods have been extensively studied for multi-site damage identification. Messina et al. [81, 82] linked the damage locations and extents with the frequency changes using a sensitivity matrix that can be obtained from the intact model. The authors set the objective function to be the multiple damage location assurance criterion (MDLAC) and found the optimal solution that maximizes the MDLAC using a traversal search strategy. Yun et al. [83] detected multi-site structural damage via setting the residual force vector as the objective function for optimization. The authors adopted a parameter subset selection method to identify the most probable damage locations. It was found that the method can diminish the parameter search space and thus is computationally efficient. Sohn and Law [47] used a Bayesian probabilistic method to detect multiple damage locations on a multi-story frame structure using estimated modal parameters. The authors used a branch-and-bound search scheme and found that this approach can save the computational cost significantly. Cao and Ouyang [84] used a modified common eigenvector analysis (CEA) to estimate the structural characteristic deflection shapes (CDS's) and derived a robust damage index comprised of local shape distortions to identify multiple damage locations.

Differing from model-based methods, data-based methods generalize damage identification to a supervised/unsupervised learning problem in the statistical pattern recognition paradigm [85, 86]. Damage localization in the data-driven scheme is pursued by a classification algorithm that evaluates the difference in damage-sensitive features between the current structural state and the initial baseline state [87]. Modal based features ex-

tracted from structural vibrations are widely used in this scheme [88]. Compared with the model-based methods, data-based methods have the advantage in computational efficiency and uncertainty handling [5]. However, there are limited studies on data-driven methods for identifying multi-site structural damage. The majority of studies in data-based structural health monitoring focus on single-site damage identification and ignore possible multi-site damage cases. Barthorpe and Worden [5] proposed a data-based approach for multi-site damage identification using data from single-site damage only. This study solved the damage detection and localization problems using a multiclass classifier based on binary support vector machines (SVMs). The features used in the classification are the transmissibility spectra extracted from the accelerance frequency response functions (FRFs). It was found that the classifier trained with only data from normal and single-damage states has the capability of recognizing multi-site damage states.

Despite the encouraging results, the study in [5] has its limitations. The challenge faced by multi-site damage identification is compromised in the experimental configuration with sparse distribution of panels and proximal installation of sensors around the potential damage locations. The coupling issue due to physical interaction and signal interference is diminished. Moreover, distinct classifiers are devised for detecting damage on each of the five panels and trained with features that are selected with the intention to optimize that classifier, which reduces the confusion during the classifier training. Real structures may not have such independence between regions of potential damage locations.

When solving engineering problems using a data-driven method through pattern recognition or machine learning, it is of great importance to properly define the problem. An appropriate definition illustrates the problem in an conceptually and structurally effi-

cient manner and helps the learning algorithms capture more meaningful information in an unambiguous way [89]. The problem of multi-site structural damage detection, in essence, is a single-instance-multi-label (SIML) classification problem, in which each object (damage case) contains various information that reflects the structural state with multiple damage locations. In comparison, MCC, method used in [5], treats each damage case, either with single or multiple damage sites, as an independent class, which ignores the inherent correlations among physically connected classes.

Zhou et al. [89] transformed the single-instance-multi-label (SIML) problem to a multi-instance-multi-label (MIML) learning problem using the *InsDif* algorithm and obtained improved classification results. *InsDif* assumes that the spatial distribution of the instances with multiple labels decodes the mixed information contained in the instances in a way beneficial for discriminating the class labels bonded to the instances. Breaking a single instance into a number of sub-instances makes the classification information more explicit, with each sub-instance corresponding to one class label. Instance differentiation, the core of the *InsDif* algorithm, transforms each data sample into a bag of instances with each reflecting the relationship of the sample with a certain possible class, which addresses the input ambiguity of multi-label classification problem. An object with multiple labels mixes the information corresponding to the different labels and makes the problem difficult for learning. The transformation from SILM to MIML problem detaches the mixed information and reduces the confusion during the learning process, and thus generates improved outcomes for this problem with input ambiguity. In data-based SHM, if the multi-site damage identification problem can be simulated as a multi-label classification task, the *InsDif* algorithm has the potential to take advantage of the information shared by cases

with common damage sites and thus can largely improve the efficiency and accuracy of identification.

To develop a more general and effective methodology, this study considers the multi-site damage identification as a multi-label classification (MLC) problem and attempts to solve it using the *InsDif* algorithm. The proposed method is examined on structural damage data from numerical simulation and experimental testing. The results are compared with that from MCC and binary classification (BC) approaches. It shows that the MLC approach outperforms the MCC and BC approaches in identifying multiple damage sites. In addition to improving the damage detection accuracy, MLC has the potential of relieving the data insufficiency problem in data-based structural health monitoring. Hence, the MLC scheme has the potential to be a promising approach for multi-site structural damage identification when using data-based methods.

4.2. Motivation and Methodology

4.2.1. Motivation

Identifying structural damage locations requires features that are sensitive to the damage location and independent of its severity. While proved efficient in localizing single-site damage, many features may not work for multi-site damage identification. The frequency change ratio has been proved independent of the damage severity but its occurring location [90], and thus it has been widely used for data-driven structural damage identification [25]. Essentially, the natural frequency of a structure is a nonlinear function of the values and spatial distribution of its stiffness and mass. To better understand the re-

lationship between frequency variation and stiffness variation, first-order Taylor series is used in this study to provide physical insight into this complex problem referring to the derivations in [90, 82]. When stiffness reduction happens at an unknown location m on the structure, the frequency changes of the i^{th} and j^{th} modes are:

$$\Delta f_i = \frac{\partial f_i}{\partial k_m} \Delta k_m \tag{4.1}$$

and

$$\Delta f_j = \frac{\partial f_j}{\partial k_m} \Delta k_m \tag{4.2}$$

in which f_i and f_j are the natural frequencies of the i^{th} and j^{th} modes; k_m denotes the structural stiffness at location m; $\frac{\partial f_i}{\partial k_m}$ is the sensitivity of f_i to k_m variation. It is noted that $\frac{\partial f_i}{\partial k_m}$ is a function of the damage location (i.e., k) and independent of its severity Δk_m. Taking the ratio between Δf_i and Δf_j yields:

$$\frac{\Delta f_i}{\Delta f_j} = \frac{\frac{\partial f_i}{\partial k_m}}{\frac{\partial f_j}{\partial k_m}} \tag{4.3}$$

which cancels out the effect of damage severity Δk_m and preserves the information about the damage location.

However, for a multi-site damage case, the equations become:

$$\Delta f_i = \frac{\partial f_i}{\partial k_1} \Delta k_1 + \frac{\partial f_i}{\partial k_2} \Delta k_2 + ... + \frac{\partial f_i}{\partial k_N} \Delta k_N, \tag{4.4}$$

$$\Delta f_j = \frac{\partial f_j}{\partial k_1} \Delta k_1 + \frac{\partial f_j}{\partial k_2} \Delta k_2 + ... + \frac{\partial f_j}{\partial k_N} \Delta k_N, \tag{4.5}$$

and

$$\frac{\Delta f_i}{\Delta f_j} = \frac{\frac{\partial f_i}{\partial k_1} \Delta k_1 + \frac{\partial f_i}{\partial k_2} \Delta k_2 + ... + \frac{\partial f_i}{\partial k_N} \Delta k_N}{\frac{\partial f_j}{\partial k_1} \Delta k_1 + \frac{\partial f_j}{\partial k_2} \Delta k_2 + ... + \frac{\partial f_j}{\partial k_N} \Delta k_N} \tag{4.6}$$

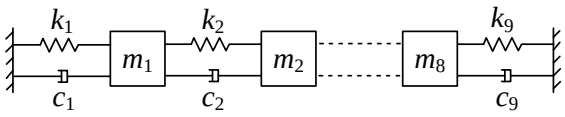


Figure 4.1. Configuration of the 8-DOF lumped mass system. m, c, and k represent mass, damping, and spring stiffness, respectively.

One can find that the ratio of frequency change cannot exclude the influence from the damage severities $\Delta k_m (m = 1, 2, ...N)$, unless they are mutually proportional, which is hardly possible in real situation. This fact invalidates the features based on frequency changes for multi-site damage identification. Similar issues can happen to many other modal-based features [91, 25]. Therefore, the identification of multi-site damage necessitates extracting appropriate features that can reflect the contribution from each potential damage location to the resulting structural responses.

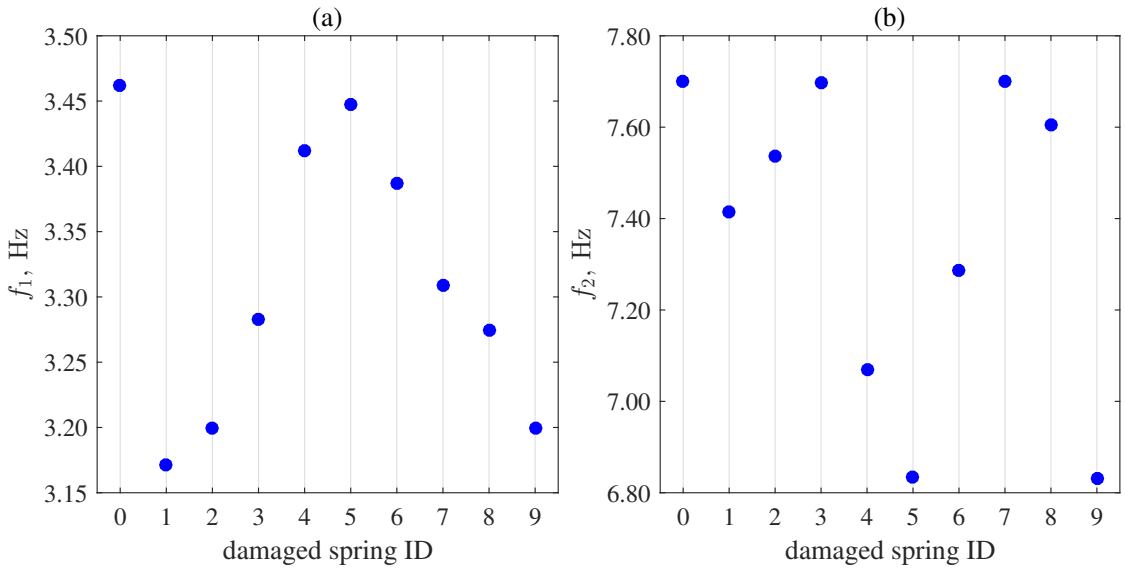


Figure 4.2. Frequency variation with single-site damage at different locations: (a) 1st mode natural frequency; (b) 2nd mode natural frequency. 0 on the x axis denotes the case with the intact structure.

To elaborate the difficulty in multi-site structural damage identification, this section uses an 8-DOF system to numerically analyze the sensitivity of vibrational properties to

Table 4.1. Mechanical properties of the 8-DOF model.

Item	I	II	III	IV	V	VI	VII	VIII	IX
m	1	2	3	4	2	2	1	3	-
c	20	20	20	20	20	20	20	20	20
k	10000	10000	10000	10000	10000	10000	10000	10000	10000

damage occurrences. The 8-DOF system contains eight lumped masses, nine springs, and nine dampers, as shown in Figure 4.1. Table 4.1 lists the non-dimensional values of the masses, spring constants, and damping coefficients that are assigned to the model.

The sensitivity analysis in this section considers cases with up to three damage locations. Damage is introduced by stiffness reduction on certain springs, and the damage severity is set as $r = 50\%$ if without specification. The sensitivity is evaluated using the first two modal frequencies, that is f_1 and f_2, considering the difficulty in obtaining high-order modal information in practical implementation. Figures 4.2 (a) and (b) depict the frequency variations with the shift of single damage locations. Both f_1 and f_2 show significant fluctuations with the shift of damage locations. The variation of f_1 demonstrates apparent symmetry because of the symmetry in the 8-DOF system. Differing from f_1, f_2 does not show continuous variation as the damage location shifts from k_1 to k_9. Additionally, the damage on k_3 and k_7 does not reduce f_2 significantly when compared with damages at other locations.

This sensitivity analysis considers two scenarios for the cases with more than one damage sites. The first scenario considers cases with exclusive damage locations, that is $[k_1, k_2]$, $[k_3, k_4]$, $[k_5, k_6]$, and $[k_7, k_8]$ for 2-site damage and $[k_1, k_2, k_3]$, $[k_4, k_5, k_6]$, and $[k_7, k_8, k_9]$ for 3-site damage; The second scenario considers cases sharing one or two damage locations, that is $[k_1, k_2]$, $[k_1, k_3]$, $[k_1, k_4]$, $[k_1, k_5]$, $[k_1, k_6]$, $[k_1, k_7]$, and $[k_1, k_8]$ for 2-

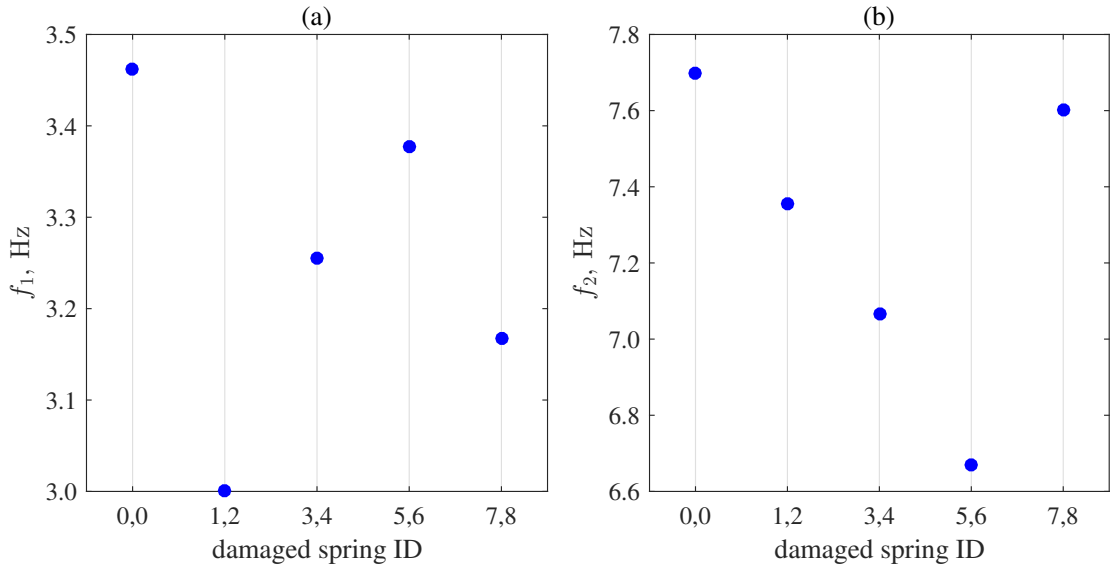


Figure 4.3. Frequency variation with double-site damage at different locations (I): (a) 1st mode natural frequency; (b) 2nd mode natural frequency. 0,0 on the x axis denotes the case with the intact structure.

site damage and $[k_1, k_2, k_3]$, $[k_1, k_2, k_4]$, $[k_1, k_2, k_5]$, $[k_1, k_2, k_6]$, $[k_1, k_2, k_7]$, $[k_1, k_2, k_8]$, and $[k_1, k_2, k_9]$ for 3-site damage.

Figures 4.3 (a) and (b) plot the variation of f_1 and f_2 for 2-site damage cases in the first scenario. As expected, additional damage further reduces the modal frequencies from the shared 1-site damage cases. The variation of f_1 still demonstrates certain symmetry with the damage location shifts. Moreover, frequencies corresponding to exclusive 2-site damage cases are sparsely distributed over the variation scales as the ambiguity caused by shared damage locations is excluded. Figures 4.4 (a) and (b) show the variations of frequencies in the second scenario. To show the difficulty of identifying multi-site damage, Figure 4.4 (a) also denotes f_1 of cases with 1-site damage on k_1 with varying damage severities ($r = 50\% - 70\%$). It shows that f_1 of 2-site damage cases at a severity of 50% can be very close to f_1 of the 1-site damage case at the shared location at a certain damage severity. Figure 4.4 (b) denotes f_2 of 1-site damage cases with 50% damage. Even with

the same damage severity, the case with damage on k_1 and k_7 is not distinguishable from that with 1-site damage on k_1 due to the limited contribution of k_7 to the variation of f_2. It can be expected that when damage severity varies, cases with multi-site damage will be much more difficult to differentiate using frequency variations.

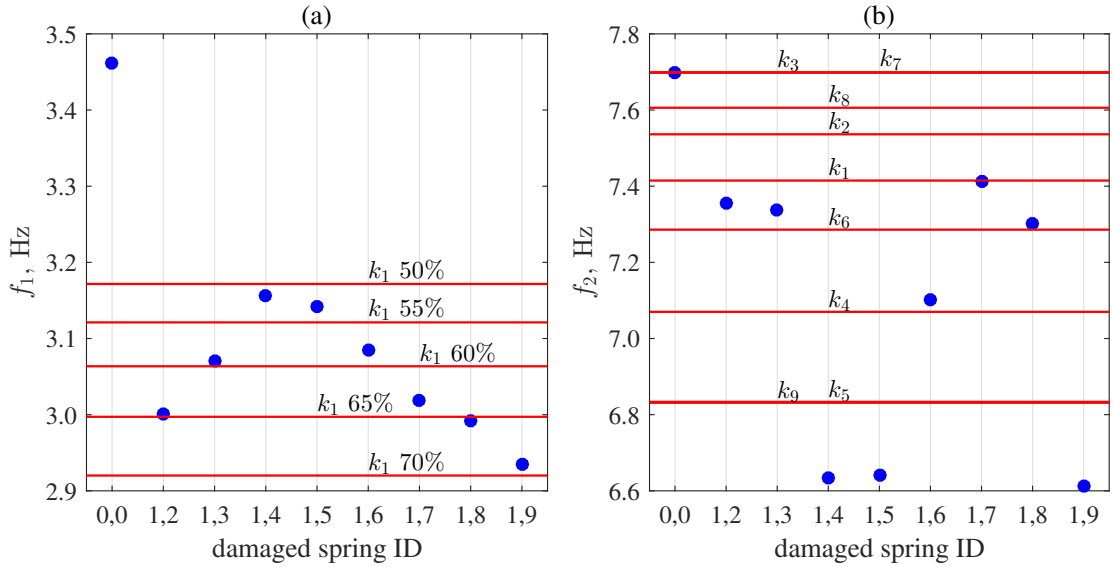


Figure 4.4. Frequency variation with double-site damage at different locations (II): (a) 1st mode natural frequency; (b) 2nd mode natural frequency. 0,0 on the x axis denotes the case with the intact structure. The reference lines denote the frequencies with 1-site damage and the percentage indicates the damage severity.

Figures 4.5 and 4.6 show the variations of frequencies for the two scenarios of 3-site damage cases. Similar to the 2-site damage cases, Figures 4.5 (a) and (b) show the sparsity of frequencies for exclusive damage cases, and Figures 4.6 (a) and (b) demonstrate the challenge of identifying multi-site damage locations using frequency variations. For example, Figure 4.6 (a) shows that f_1 of the damage case $[k_1, k_2, k_3]$ with a severity of $r = 50\%$ almost coincides with that of the case $[k_1, k_2]$ with $r = 55\%$; in Figure 4.6 (b), The case $[k_3, k_4]$ has very close magnitude of f_2 to the case $[k_1, k_2, k_6]$.

In summary, this section highlights the motivation of this study by demonstrat-

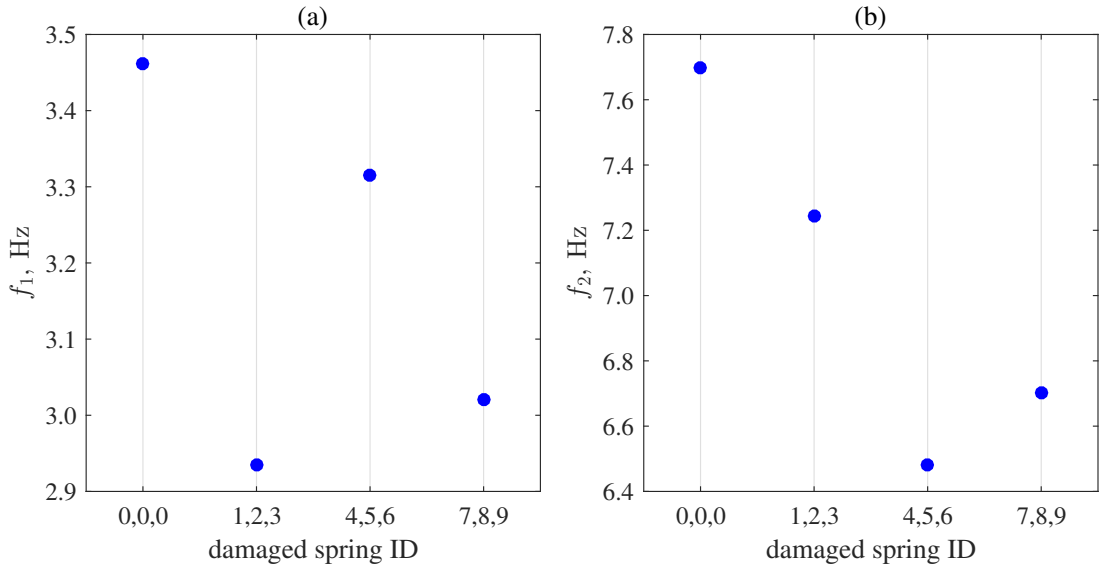


Figure 4.5. Frequency variation with 3-site damage at different locations (I): (a) 1st mode natural frequency; (b) 2nd mode natural frequency. 0,0,0 on the x axis denotes the case with the intact structure.

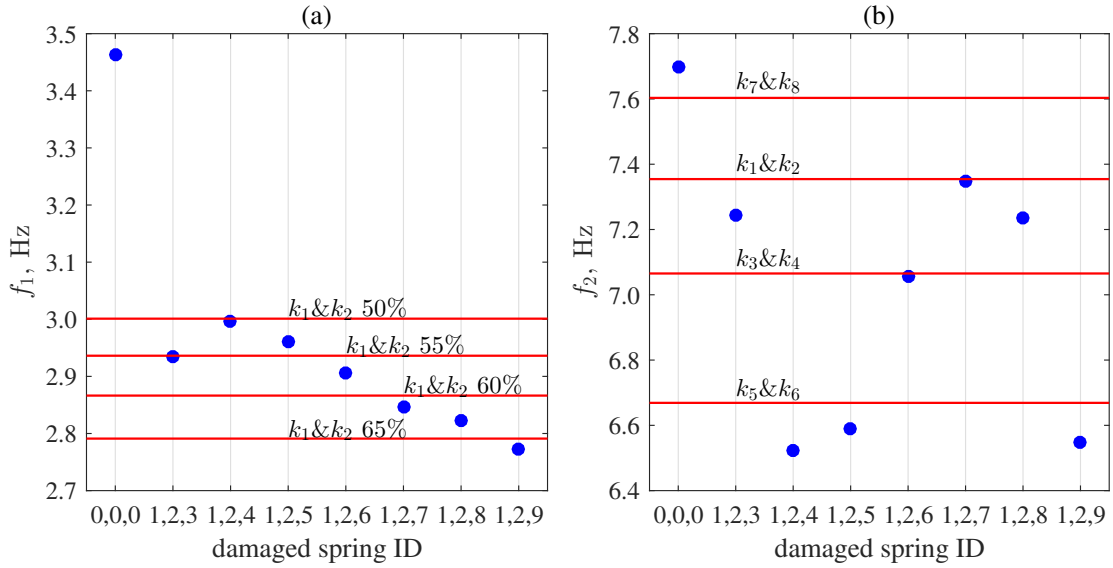


Figure 4.6. Frequency variation with 3-site damage at different locations (II): (a) 1st mode natural frequency; (b) 2nd mode natural frequency. 0,0,0 on the x axis denotes the case with the intact structure. The reference lines denote the frequencies with 2-site damage and the percentage indicates the damage severity.

ing the challenge of identifying multi-site structural damage with vibration-based meth-

ods. Theoretical analysis and numerical simulation approves that multi-site damage occur-

rence introduce ambiguity to the influence of damage on vibrational properties especially when considering the variation of damage severities. Hence, an advanced method of pattern recognition is necessary to resolve the ambiguity contained in multi-site damage cases.

4.2.2. Methodology

This study proposes identifying multi-site structural damage using a data-based method with the MLC scheme of machine learning. In this study, the multi-label learning denote each damage case with a label vector in which each element shows a damage occurrence site. Let $\mathcal{X} = \mathbb{R}^d$ denote the input space of extracted damage sensitive features, and $\mathcal{Y} = \{1, 2, ..., Q\}$ denote the set of possible labels, each of them representing one damage location. Q represents the total number of possible damage locations considered in a certain case. The training data set of MLC can be formulated as $S = \{(x_1, Y_1), (x_2, Y_2), ..., (x_N, Y_N)\}$, in which $x_i \in \mathcal{X}$ is an instance standing for the feature vector of a damage state and $Y_i \subseteq \mathcal{Y}$ is a set of labels associated with x_i denoting its damage sites. Then MLC aims to learn a function $h : \mathcal{X} \rightarrow 2^{\mathcal{Y}}$ from S that can be used to identify multi-site damage including damage states which are not included in the training data set.

Figure 4.7 illustrates the proposed framework of multi-site structural damage identification using the MLC method. The framework consists of three modules: data preparation, instance differentiation, and MIML classifier. Under external excitation, the structural response $\mathbf{R}$ containing the displacement $\mathbf{u}$, velocity $\dot{\mathbf{u}}$, and acceleration $\ddot{\mathbf{u}}$ will be obtained. The responses $\mathbf{R}$ are subsequently used to extract modal properties, including the natural frequencies $\mathbf{f}$, mode shapes $\mathbf{\Phi}$, and mode shape curvatures $\mathbf{\Phi}''$, with a blind source separation method introduced in [92]. Feature extraction and selection using the modal

properties yield features sensitive to structural damage occurrence. These procedures constitute the data preparation module that prepares the feature vectors of all the damage cases for subsequent data transformation using the *InsDif* algorithm.

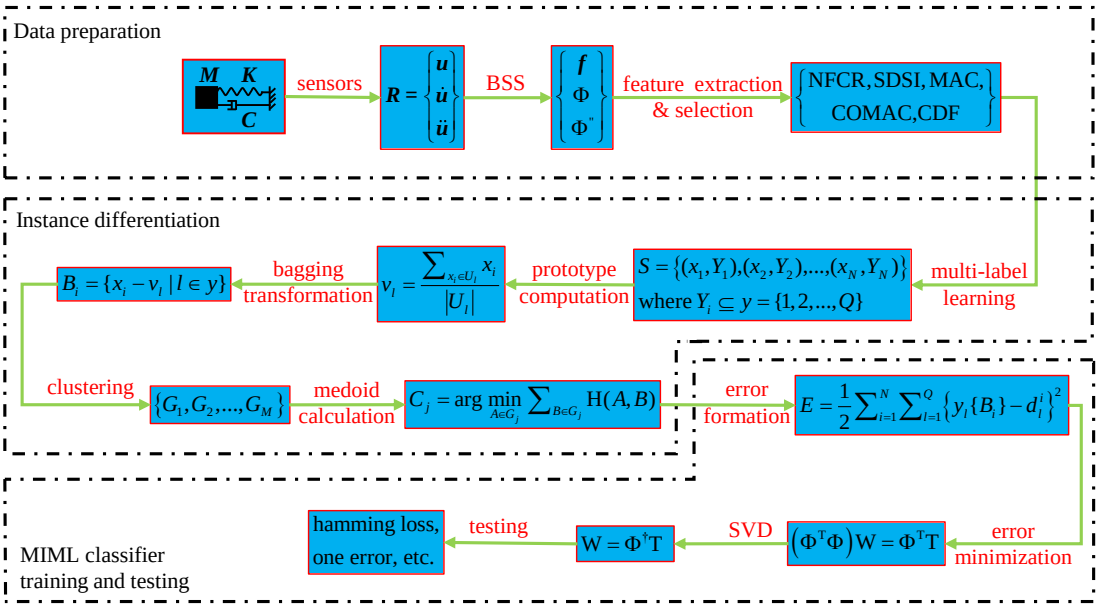


Figure 4.7. Framework of multi-site damage identification using multi-label classification

Given the feature vectors and corresponding damage states of all available cases, the multi-label learning problem can be formulated with the partition of training data set, S. To implement MLC, *InsDif* first calculates the prototype vector v_l for each class label c_l by averaging all the training instances labeled with c_l:

$$v_l = \frac{\sum_{x_l \in U_l} x_i}{|U_l|}, \text{ in which } U_l = \{x_i | \{x_i, Y_i\} \in S, l \in Y_i\}, l \in \mathcal{Y} \tag{4.7}$$

in which $|U_l|$ denote the quantity of instances in U_l. The prototype vector v_l describes the common characteristics shared by instances labeled with c_l. After obtaining the prototype vectors, *InsDif* then represents each example x_i with a bag of instances $\mathbf{B}_i$:

$$\mathbf{B}_i = \{x_i - v_l | l \in \mathcal{Y}\} \tag{4.8}$$

Hence, *InsDif* transforms each example x_i into a bag $\mathbf{B}_i$ that has a dimension equaling the number of possible classes, Q. Then the training set is transformed to $S^{\text{new}} = \{(\mathbf{B}_1, Y_1), (\mathbf{B}_2, Y_2), ..., (\mathbf{B}_N, Y_N)\}$. Thus, the MLC problem is transformed from a SIML learning problem to a MIML learning problem. This transformation attempts to exploit the spatial difference between x_i and v_l for improved learning performance.

To formulate the classification problem, *InsDif* clusters the transformed training bags $\mathbf{B}_i, (i = 1, 2, ..., N)$ into M disjoint groups of bags $\mathbf{G}_j, (j = 1, 2, ..., M)$ using the k-medoid algorithm:

$$\bigcup_{j=1}^{M} \mathbf{G}_j = \{\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_N\}, \text{ and } \mathbf{G}_i \bigcap_{i \neq j} \mathbf{G}_j = \emptyset \tag{4.9}$$

in which M is a user specified parameter that does not affect much the final classification results. In the k-medoid clustering, the distance between two instance bags $A = \{a_1, a_2, ..., a_{n1}\}$ and $B = \{b_1, b_2, ..., b_{n2}\}$ are measured by the Hausdorff distance, $\mathrm{H}(A, B)$:

$$\mathrm{H}(A, B) = \max\left\{\max_{a \in A} \min_{b \in B} \|a - b\|, \max_{b \in B} \min_{a \in A} \|b - a\|\right\} \tag{4.10}$$

in which $\|a - b\|$ denotes the Euclidean distance between instances a and b. Then the medoids $\mathbf{C}_j \, (1 \leqslant j \leqslant M)$ of the$M$ partitions can be computed as

$$\mathbf{C}_j = \arg\min_{A \in G_j} \sum_{B \in G_j} \mathrm{H}(A, B) \tag{4.11}$$

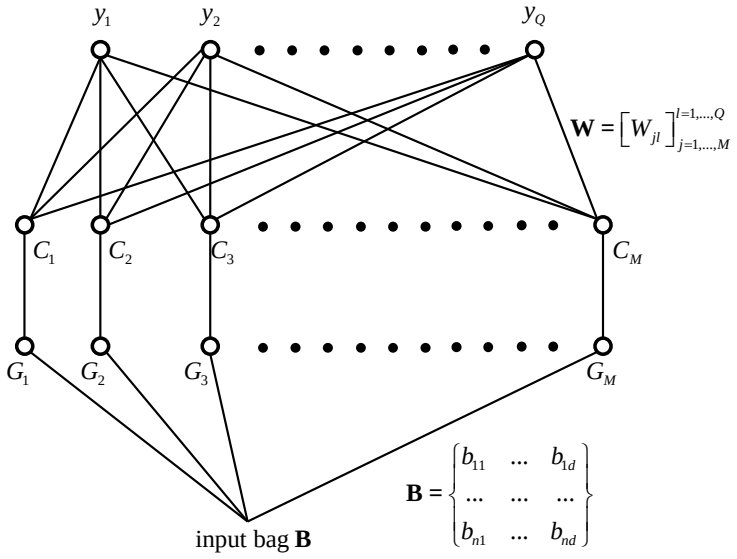


Figure 4.8. Two-level classification structure used by *InsDif* (reproduced from [89])

Clustering has the potential to discover the underlying information of a data set that contains confusing ambiguity superficially. Hence, the medoid of each partition may contain important information linking the partitioned bags to their multiple class labels. After this clustering partition, the instance bags B are converted to an M-dimensional feature vector $[\phi_1(B), \phi_1(B), ..., \phi_M(B)]^{\mathrm{T}}$, in which $\phi_j(B) = \mathrm{H}(B, C_j)$. Then the output of the second level, i.e. the weights $\mathbf{W} = [w_{jl}]_{M \times Q}$, is obtained by minimizing the following error function:

$$E = \frac{1}{2} \sum_{i=1}^{N} \sum_{l=1}^{Q} \left\{ y_l(B_i) - d_l^i \right\}^2 \tag{4.12}$$

in which $y_l(B_i) = \sum_{j=1}^{M} w_{jl} \phi_j(B_i)$ is the current output of B_i on the l^{th} class using the assumed weights, and d_l^i is the target output that equals +1 if $l \in Y_i$ and -1 if $l \notin Y_i$. Differentiating the objective error function in Equation 4.12 with respect to w_{jl} and setting the derivative to zero yields the following normal equations:

$$\left(\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi}\right)\mathbf{W} = \mathbf{\Phi}^{\mathrm{T}}\mathbf{T} \tag{4.13}$$

in which $\mathbf{\Phi} = [\phi_{ij}]_{N\times M}$ with $\phi_{ij} = \phi_j(B_i)$ and $\mathbf{T} = [d_l^i]_{N\times Q}$. This equation can be solved using the singular value decomposition (SVD), such that

$$\mathbf{W} = \mathbf{\Phi}^{\dagger}\mathbf{T} \tag{4.14}$$

in which $\mathbf{\Phi}^{\dagger}$ is the pseudo-inverse of $\mathbf{\Phi}$.

Figure 4.8 illustrates the two-level MIML classification using the *InsDif* algorithm. Finally, using the computed classification weights $\mathbf{W}$ obtained with the training data set, the concluding procedure of this framework evaluates the classification quality using the test data set. The evaluation indices include hamming loss, one error, coverage, ranking loss, and average precision. Additionally, the classifier outputs should be analyzed to evaluate the classification quality in more detail.

4.3. Feature Extraction

To achieve efficient structural damage identification and evaluation using machine learning methods, the extraction of damage-sensitive features that are capable of reflecting the damage location and severity is of critical importance. Extracting and selecting features that can distinguish damaged structure and varied damage locations and severities is one primary component of structural health monitoring [93]. Indicative features of structural damage lead to accurate and efficient damage diagnosis, while irrelevant features results in misleading judgment on damage locations and severities. The following subsections

present a review on some features that are widely used for structural health monitoring.

4.3.1. Natural Frequency Features

Adams et al. initiated damage detection using frequency variations in 1970s [94], and it was extended by Crawly and Adams [90]. The global nature of natural frequency makes it not an superior candidate for damage localization [95, 96, 97]. Nevertheless, it has been verified that the frequency change ratio of two modes caused by damage is not a function of the damage extent but its occurrence location [90]. In addition, the normalized frequency change ratio (NFCR) is also proved to be related to the damage location and independent from the damage severity [25] . The fractional frequency change (FFC) for the i^{th} mode is defined as

$$\text{FFC}_i = \frac{f_{\text{u}i} - f_{\text{d}i}}{f_{\text{u}i}} \tag{4.15}$$

in which $f_{\text{u}i}$ and $f_{\text{d}i}$ are the i^{th} mode frequencies of the structure in intact and damaged states, respectively. The NFCR for the i^{th} mode is defined as:

$$\text{NFCR}_i = \frac{\text{FFC}_i}{\sum_{j=1}^{N} \text{FFC}_j} \tag{4.16}$$

in which N is the number of modes with measured natural frequencies.

4.3.2. Mode Shape Features

It is worth noting that all the frequency based features cannot discriminate between two symmetric damage locations. In comparison, mode shapes contain structural spatial information that is valuable for damage localization [98] and are less susceptible to environmental influence [99].

(1) ***DSI****(Damage Signature Index)* A combined damage signature index (**DSI**) consisting of both mode shapes and frequencies were proposed in [91]. It was found that the

index **DSI** is dependent on damage location only. The **DSI** is defined as the deformation of mode shape vector normalized by the modal frequency square variation, that is,

$$\mathbf{DSI}_i = \frac{\mathbf{\Phi}_{\mathrm{u}i} - \mathbf{\Phi}_{\mathrm{d}i}}{|f_{\mathrm{u}i}^2 - f_{\mathrm{d}i}^2|} \tag{4.17}$$

in which $\mathbf{\Phi}_{\mathrm{u}i}$ and $\mathbf{\Phi}_{\mathrm{d}i}$ are the modal shape vectors of the i^{th} mode in undamaged and damaged conditions, respectively.

(2) ***SDSI****(Simplified Damage Signature Index)* Instead of a scalar parameter, **DSI** is series of vectors, with its quantity equal to the number of selected modes and its dimension equal to the number of degree of freedoms (DOFs) for mode shape extraction. For simplicity, the chapter simplifies this mode shape feature to an index that averages the contribution of each mode for a certain DOF, following the practice of mode shape curvature feature that will be introduced later. This simplification generates the simplified damage signature index (**SDSI**) as

$$\mathbf{SDSI} = \frac{1}{N}\sum_{i=1}^{N} |\mathbf{DSI}_i| = \frac{1}{N}\sum_{i=1}^{N} \frac{|\mathbf{\Phi}_{\mathrm{u}i} - \mathbf{\Phi}_{\mathrm{d}i}|}{|f_{\mathrm{u}i}^2 - f_{\mathrm{d}i}^2|} \tag{4.18}$$

in which N is the number of selected modes for damage detection.

(3) ***MAC*** *(Modal Assurance Criterion)* Modal Assurance Criterion (**MAC**), proposed by Allemang and Brown [100] detects mode shape shifts taking advantage of the orthogonality of mode shapes. **MAC** for the i^{th} mode shape is defined as

$$\mathbf{MAC}_i = \frac{\left|\mathbf{\Phi}_{\mathrm{u}i}^{\mathrm{T}}\mathbf{\Phi}_{\mathrm{d}i}\right|^2}{\left(\mathbf{\Phi}_{\mathrm{u}i}^{\mathrm{T}}\mathbf{\Phi}_{\mathrm{u}i}\right)\left(\mathbf{\Phi}_{\mathrm{d}i}^{\mathrm{T}}\mathbf{\Phi}_{\mathrm{d}i}\right)} \tag{4.19}$$

in which $\mathbf{\Phi}_{\mathrm{u}i}$ is the i^{th} mode shape of the undamaged structure and $\mathbf{\Phi}_{\mathrm{d}i}$ is that of the damaged structure. **MAC** is a metric that measures the correlation level of mode shapes of a structure before and after damage. The value of **MAC** ranges between 0 and 1. A

value of 1 indicates 100% vector correlation and no obvious mode shape variation after damage occurrence; a value of 0 denotes no correlation, and it means significant mode shifts due to the damage occurrence. Zhao and Zhang proved that **MAC**s of certain modes are sensitive to structural damage using an example of truss bridge [101].

(4) ***COMAC*** *(Coordinate Modal Assurance Criterion)* Kim extended **MAC** and developed the coordinate modal assurance criterion (**COMAC**) that identifies the coordinates where the mode shapes do not agree before and after structural damage occurrence [102].

$$\mathbf{COMAC}_i = \frac{\left[\sum_{j=1}^{N} \mathbf{\Phi}_{\mathrm{u},j}^{i} \mathbf{\Phi}_{\mathrm{d},j}^{i}\right]^2}{\sum_{j=1}^{N} \left(\mathbf{\Phi}_{\mathrm{u},j}^{i}\right)^2 \sum_{j=1}^{N} \left(\mathbf{\Phi}_{\mathrm{d},j}^{i}\right)^2} \tag{4.20}$$

in which the subscript i stands for the i^{th} DOF of mode shape, j the mode order, and N the total number of modes taken into consideration.

4.3.3. Mode Shape Curvature Features

Pandey et al. [103] for the first time proposed to use the modal curvature for damage detection. The modal curvature of the j^{th} mode at DOF i can be calculated using Equation 4.21.

$$\mathbf{\Phi}_j^{''i} = \frac{\mathbf{\Phi}_j^{i+1} - 2\mathbf{\Phi}_j^{i} + \mathbf{\Phi}_j^{i-1}}{h^2} \tag{4.21}$$

in which $\mathbf{\Phi}_j^i$ is the j^{th} mode shape amplitude at the i^{th} DOF; h is the distance between two successive measured locations.

Abdel et al. [104] used the modal curvature to detect bridge damage and obtained promising results. However, modal curvature works well for lower mode shape while higher modal curvature might produce false damage indication. Therefore, features extracted

Table 4.2. Results of feature extraction

feature ID	feature content
1	NFCR of 1st mode
2	NFCR of 2nd mode
3-10	SDSI of Mass 1-8
11	MAC of 1st mode
12	MAC of 2nd mode
13-20	COMAC of Mass 1-8
21-26	CDF of Mass 2-7

from mode shape curvatures should be used with caution.

(1) ***CDF****(Curvature Damage Factor)* Wahab and Roeck proposed the curvature damage factor that summarizes the contribution of curvature from all the modes [104], as

$$\mathbf{CDF} = \frac{1}{N}\sum_{j=1}^{N}\left|\mathbf{\Phi}''_{\mathrm{u}j} - \mathbf{\Phi}''_{\mathrm{d}j}\right| \tag{4.22}$$

in which N is the total number of modes considered in the analysis, $\mathbf{\Phi}''_{\mathrm{u}j}$ is the j^{th} mode shape curvature of the intact structure, and $\mathbf{\Phi}''_{\mathrm{d}j}$ is that of the damaged structure.

Table 4.2 lists the result of feature extraction using the modal properties. Again, the present study uses the first two modal information for feature extraction considering the difficulty of obtaining higher-frequency modes in practice. For the features from mode shapes and mode shape curvatures, the responses of the 8 masses are used for modal information extraction. As a result, **COMAC** has a dimension of 8 and **CDF** 6. The extracted features will be used for structural damage identification using machine learning methods in the rest of this chapter.

4.4. Numerical Study: Multiclass, Binary, and Multi-Label Classification

This section numerically studies the multi-site damage identification problem using the 8-DOF system shown in Figure 4.1. Damage is introduced to one to three locations through reducing the stiffness of the corresponding springs by a random percentage from 10% to 90%. A wide range of stiffness reduction ratios is applied to consider the varied damage severities in real engineering. In this study, a total number of 129 damage cases are considered: 9 ($\binom{9}{1}$) for 1-site damage, 36 ($\binom{9}{2}$) for 2-site damage, and 84 ($\binom{9}{3}$) for 3-site damage. For each scenario with the introduced damage, an impact loading is applied to mass m_4, and the dynamic responses, including the acceleration, velocity, and displacement of the eight masses, are collected. This study uses a blind source separation (BSS) method based on independent component analysis (ICA) [92] to extract the modal properties from the measured responses. Then the features listed in Table 4.2 are derived following the procedures described in Section 4.3.

This section uses three aforementioned machine learning methods, MCC, BC, and MLC for multi-site damage identification and evaluates their performance. Based on comparison, this study uses support vector machine (SVM) classifiers for MCC and artificial neural network (ANN) classifiers for BC. Radial basis function (RBF) kernel is used in SVM. ANN in BC uses a 3-hidden multi-layer perceptron (MLP) model with the number of neurons set as [50 30 50]. The *InsDif* algorithm introduced in Section 4.2.2 is used to implement the MLC. Hyperparameters of the used classification models are tuned in model validation, and only testing results are presented in the case studies presented in this chapter.

It should be noted that the comparison of the classification performance of the three schemes is not straightforward. MCC quality is assessed based on the labeling accuracy of the testing examples belonging to each class, whether corresponding to a single-site or multi-site damage case. It does not evaluate the accuracy of identifying each damage location in a multi-site damage case. In contrast, MLC evaluates the classification quality of each label for an example case with multi-site damage. Differing from MCC and MLC, BC yields a series of classifiers, with each evaluating the damage condition of a testing example regarding one possible damage site. Hence, the output of BC can be integrated with that of MLC by summarizing the labeling quality of an example regrading the damage condition of each site, and thus are comparable to that of MLC.

MCC quality can be evaluated using three metrics, including precision, recall, and F-score that will be introduced as follows. In a classification problem, the precision for a certain category is defined as the ratio between the number of true positives and the sum of quantity of true positives and false positives. True positives are the instances that are correctly classified into that category, while false positives are the instances that are wrongly labeled as belonging to that category. Recall is a ratio between the number of true positives and the total number of instances that indeed belong to the corresponding category. Equations (4.23) and (4.24) express the formula of precision and recall, respectively [105], in which N_{tp}, N_{fp}, and N_{fn} denote the number of true positives, false positives, and false negatives, respectively. A precision of 1.0 for a category indicates that all the instances that are classified into that category are indeed instances of that class, while a recall of 1.0 means that all the instances in that category have been classified correctly. F-score is a metric of classification quality that combines the concepts of precision and recall

[105]. As can be seen from Equation 4.25, the value of F-score equals 1.0 only when both precision and recall have values of 1.0.

$$\text{Precision} = \frac{N_{\text{tp}}}{N_{\text{tp}} + N_{\text{fp}}} \tag{4.23}$$

$$\text{Recall} = \frac{N_{\text{tp}}}{N_{\text{tp}} + N_{\text{fn}}} \tag{4.24}$$

$$\text{F} - \text{score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.25}$$

MLC performance is evaluated using metrics that are different from MCC. Hamming loss is defined as the portion of instances i with labels l having the output hypothesis $\text{f}(x_i, l)$ differing from the targets $Y_i[l]$. One-error evaluates how many times none of the predicted instance labels lies in the target class. Coverage measures the average number of instances required to cover all the target labels. Ranking loss measures the average portion of instance labels that are ranked incorrectly. Average precision evaluates the average fraction of instance labels with the correct prediction [106].

BC uses the same metrics as MCC, i.e., precision, recall, and F-score, to evaluate its performance of labeling the examples to each of the target categories. As the BC results can be integrated to form a label matrix of pseudo-MLC, and thereafter can be evaluated using the metrics for MLC quality assessment.

4.4.1. Multi-Class Classification Result

MCC has been widely used for damage detection and localization by taking the intact and various damage cases (with different number and locations of damage) as the tar-

get categories for classification. It marks the cases with multiple damage locations using exclusive category labels, which neglects the interconnection between two damage cases that share some damage locations. For example, the case with three damage locations A, B, and C and that with two damage locations A and B should share some common characteristics that help detect the common damage locations (A and B) and discern the distinct damage location (C). As will be demonstrated in the examples below, MCC will require more samples to train a classifier with satisfactory performance because of the neglect of the connection between damage cases.

Figures 4.9 (a) and (b) show the MCC quality evaluated using F-scores for scenarios with 10 and 100 samples in each class (NOS), respectively. The number of damage sites (NDS) ranges from 1 to 3. A ten-fold cross validation is used to maximize the generality of the trained model with the available data. Generally, one can find in Figures 4.9 (a) and (b) that, the classification quality is better for cases with fewer NDS than for cases with more NDS. Figure 4.9 (a) shows that in the scenario with NOS = 10, the F-score is not available for certain cases because of the insufficient generality of the trained classifier. When more samples are generated for each class, the trained classifier has improved generality and thus yields reasonable and increased classification scores. The difference in classification quality among the three damage scenarios indicates that, the inter-correlation and ambiguity between cases with multiple damage sites causes confusion to the classifier training and thus reduces the classifier's generality. Moreover, the accuracy of testing the trained MCC classifier relies on the availability of tested examples with identical labels in the training dataset; or in other words, MCC cannot correctly label the examples belonging to a class that is not contained when training the classifier. This drawback further

limits the application of MCC for structural damage identification. Therefore, advanced classification techniques are needed to reveal the interconnection between damage cases with multi-site damage and resolve the confusion issue in classification.

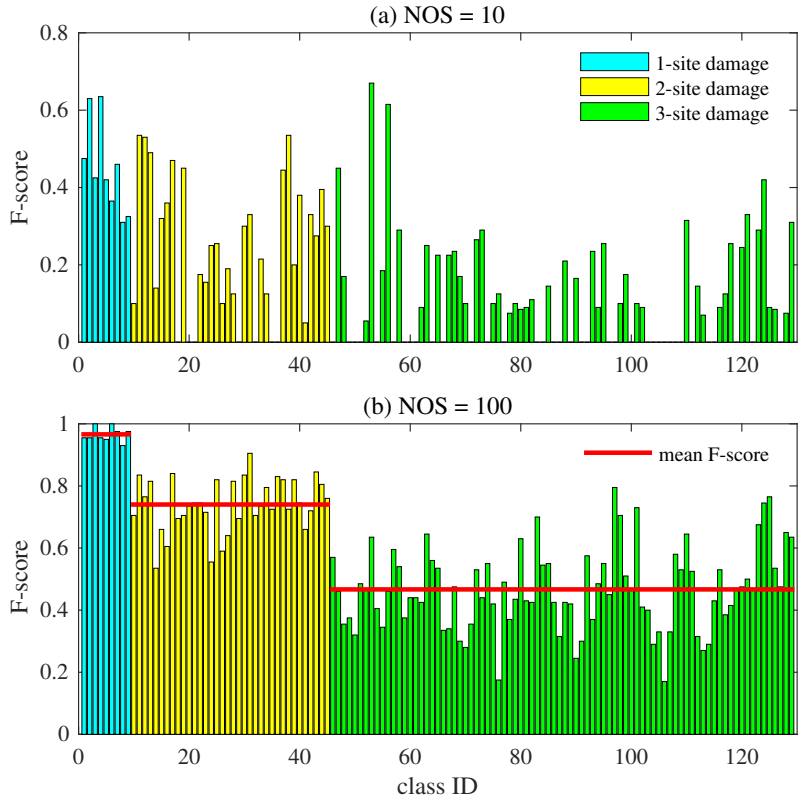


Figure 4.9. F-score of MCC for multi-site damage identification. (a) 10 samples are used in each class; (b) 100 samples are used in each class. The red reference line in each subplot denotes the mean F-score for cases with the same number of damage sites. NOS stands for number of samples in each class.

4.4.2. Binary Classification Result

BC classifies the elements of a given set into two groups using a classification rule. It can be used for damage detection with respect to a certain location of a structure, by making a decision about whether a structure possesses certain properties or characteristics. BC reduces the MLC problem, as it decomposes the classification task into multiple independent problems with binary category labels [107, 108].

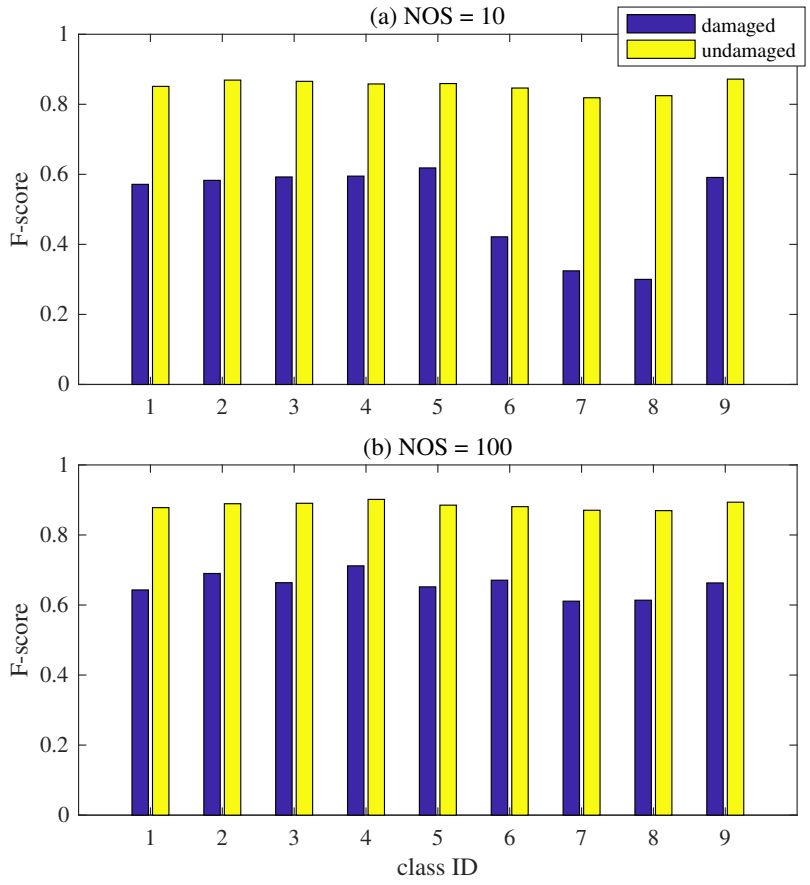


Figure 4.10. F-score of binary classification. The class IDs correspond to the possible damage locations, i.e., k_1 to k_9. NOS stands for number of samples in each class.

For damage detection and localization, BC trains all the instances for each possible damage location separately, setting the damage condition at that location (damaged or intact) as the target label, and finally integrates the classification results for all the damage locations to formulate a model for MLC decision-making. However, BC, by separating classes with multiple labels, fail to take into account the correlation between the instance's labels and thus weakens the system's expressive power [106].

Figures 4.10 (a) and (b) show the F-score of BC regarding each potential damage location for models trained with NOS = 10 and 100, respectively. It shows that BC has better classification quality for examples labeled as "undamaged" than for those la-

beled as "damaged". This difference can be attributed to the following two reasons. First, the examples used in this study for training contain damage cases with up to three damage locations. As a result, for a certain location, majority of the examples are labeled as "undamaged". In detail, the 129 damage cases with 1-3 damage locations contain only 37 cases with damage on a certain location. The second reason is that, for a specific location to be labeled as "damaged" with significant certainty in BC, the damage at that location must generate characteristics that are intense or distinguishable enough to prevent from being overwhelmed by those generated from damage elsewhere. Due to this difference in classification quality for "damaged" and "undamaged" categories, decision-making using a trained BC model is inclined to false-negative errors with "damaged" regarded as positive. This inclination is not conservative in decision-making for structural damage identification that is in favor of structural safety. Moreover, comparing the scores in Figures 4.10 (a) and (b), it shows that the increase of NOS significantly improves the classification quality on "damaged" labels while it has little effect on the "undamaged" labels.

Table 4.3 compares the MLC scores calculated using the BC outputs for cases with NDS = 1 to 3 and NOS = 100. The arrows in the first column denote the trend for better scores. Hamming loss and average precision are two quantitative metrics evaluating the average labeling quality of an MLC model. As most labels are "undamaged" that have high classification accuracy, the average precision may not reflect the classification quality objectively. In comparison, hamming loss measures the average labeling error per label and thus excludes the influence of label quantity imbalance. Table 4.3 shows that though the average precision does not change much with the NDS, the hamming loss show large increase. Additionally, the coverage and ranking loss indicates the decrease of classification

quality as the NDS increases.

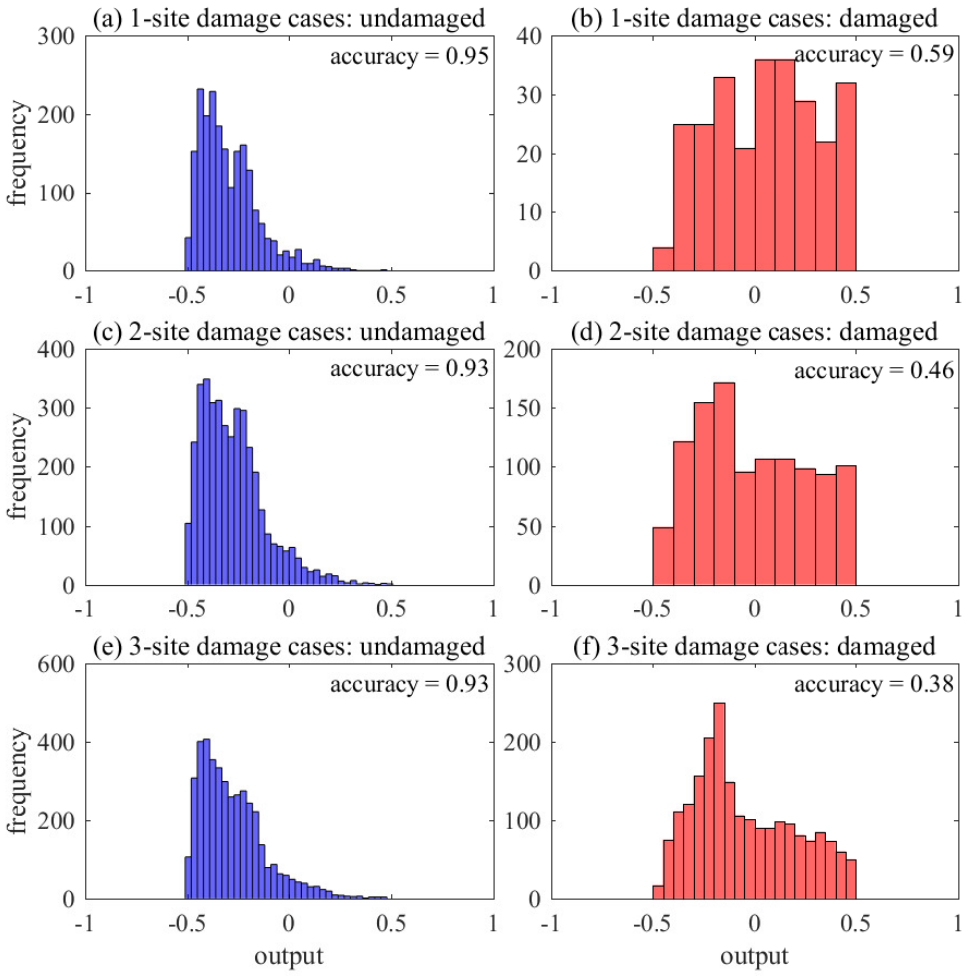


Figure 4.11. Histograms of BC outputs for cases with NDS = 1 to 3. Negative outputs predict labels as "undamaged", and positive outputs denote "damaged". Accuracy is calculated as the ratio between the number of correctly categorized labels and the total number of labels belonging to that category.

Figures 4.11 (a) to (f) show the histograms of BC outputs for "undamaged" and "damaged" labels of the cases with NOS = 1 to 3. The outputs of an MCC or BC classifier range between 0 and 1, which are shifted to the range [-0.5,0.5] for consistence with the MLC outputs regarding the sign. The classification accuracy is shown in each subplot. It shows that the classification quality for the "undamaged" labels almost remains the same as the NDS increases because of the increased sample quantity that guarantees the

trained model's generality. In contrast, the accuracy of "damaged" labels reduces largely with the increase of the NDS. Moreover, the classification accuracy of "damaged" labels remains much lower than that of "undamaged" labels. Hence, BC, as a reduced MLC approach, does not improve the performance of data-driven multi-site damage identification, mainly because it fails to resolve the ambiguity between damage cases and exploit their correlations.

Table 4.3. MLC scores calculated from BC outputs.

evaluation metrics	NDS		
	1	2	3
hamming loss ↓	0.09	0.17	0.25
one error ↓	0.37	0.29	0.21
coverage ↓	1.04	3.26	4.58
ranking loss ↓	0.13	0.20	0.22
average precision ↑	0.75	0.71	0.75

4.4.3. Multi-Label Classification Result

Multi-label classification solves the multi-site damage identification problem considering the interconnection of instances sharing common damage locations. This study uses the *InsDif* algorithm [109] to solve the multi-label classification problem using the MIML approach that separates the mixed information contained within the instances via instance differentiation and clustering. Section 4.2.2 introduces the details of this *InsDif*-MLC scheme. This section presents the results of multi-site damage identification using this method.

Before presenting the MLC results, this section first compares the feature distribution before and after applying the *InsDif* algorithm to demonstrate the advantage of *InsDif*-MLC in resolving the ambiguity lying in multi-site damage cases. Three damage

Table 4.4. MLC scores using the *InsDif* algorithm.

evaluation criterion	NDS			
	1	2	3	3*
hamming loss ↓	0.05	0.10	0.18	0.18
one error ↓	0.16	0.08	0.05	0.03
coverage ↓	0.48	2.19	3.99	4.21
ranking loss ↓	0.06	0.09	0.15	0.15
average precision ↑	0.89	0.86	0.84	0.85

* testing on unobserved 3-site damage cases when training the classifier.

cases are used for comparison, including single damage on k_1 (SD1), single damage on k_9 (SD9), and double damage on k_1 and k_9 (DD1&9). Two frequency-based features, i.e., NFCR of the first mode ($NFCR_1$) and the second mode ($NFCR_2$), are used for this analysis. NOS is set as 100 for each damage case. Figure 4.12 (a) scatters the raw features of the three cases. It shows that without the interference from the 2-site damage case, the two 1-site damage cases, SD1 and SD9, can be roughly separated by a smooth curve in the 2D feature space. However, the existence of the 2-site damage case, DD1&9, complicates the distribution of samples in the feature space. As a result, It becomes very difficult to distinguish the samples of DD1&9 from that of SD1 and SD9, which will impair the classification quality when these data are used for training a multiclass classifier.

Figure 4.12 (b) illustrates the three principal components (PC 1, PC 2, and PC 3) of the transformed features using the *InsDif* algorithm. Figure 4.12 (c) shows the data on the right half of Figure 4.12 (b) for a clearer presentation. Comparison between Figures 4.12 (a) and (b)/(c) shows that after the feature transformation, most samples of the three cases are approximately distributed on three parallel planes in the 3D space, which largely improves the separability of damage cases. Hence, it can be expected that a classifier trained with the transformed features can yield improved classification performance.

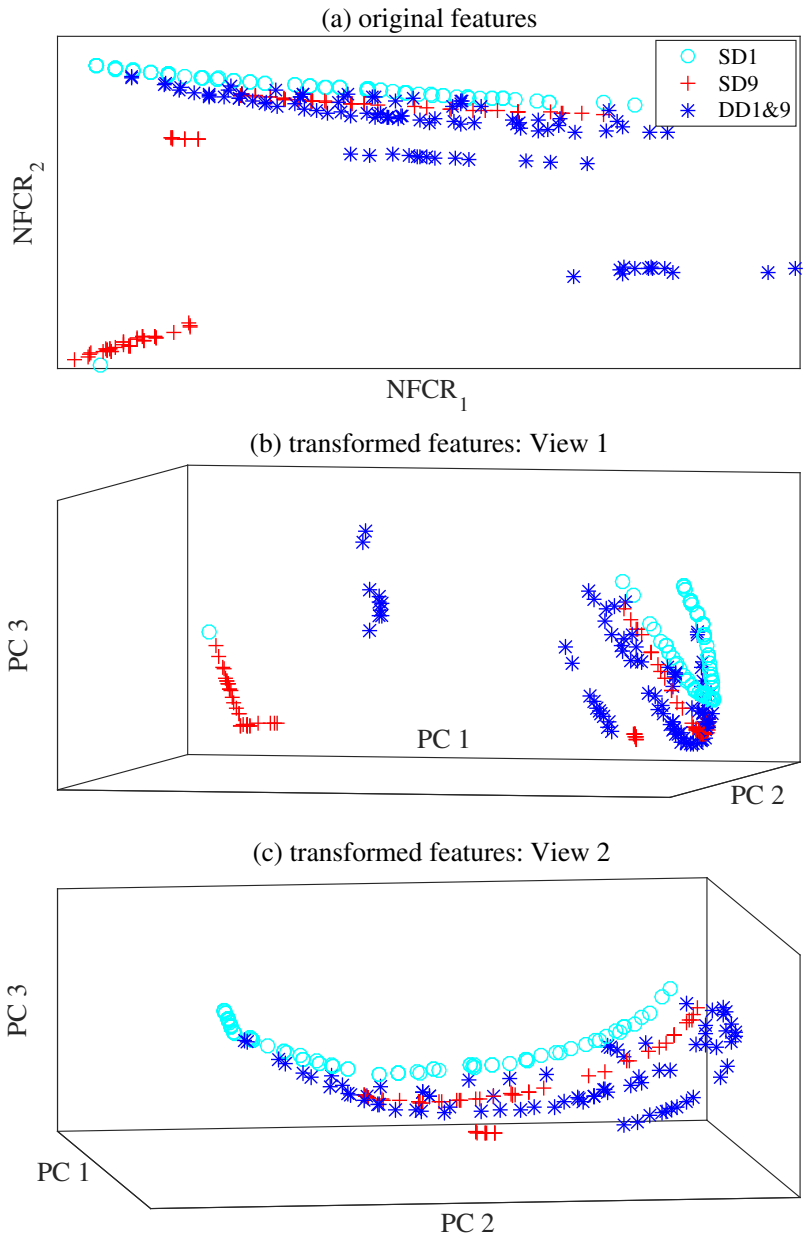


Figure 4.12. Scatter-plots of raw features and transformed features. (a) shows the scatter plot of the instances in the 2D feature space before transformation; (b) and (c) plot the instances in the transformed feature space from two different perspectives with the feature space reduced to 3D via PCA.

Table 4.4 tabulates the results of the MLC for damage localization with NOS = 100. In comparison with the results calculated from BC outputs in Table 4.3, MLC yields better scores regarding all the five metrics. As mentioned above, MCC does not support

labeling the unobserved examples when training the classifier. The fifth column of Table 4.4 lists the scores of MLC when testing on unobserved examples with NDS = 3. These scores are obtained by setting the data from using the first 120 damage cases for training and the remaining 9 cases for testing. Comparison between the fourth and fifth column shows that consistently satisfactory results can be obtained when using MLC on unobserved examples.

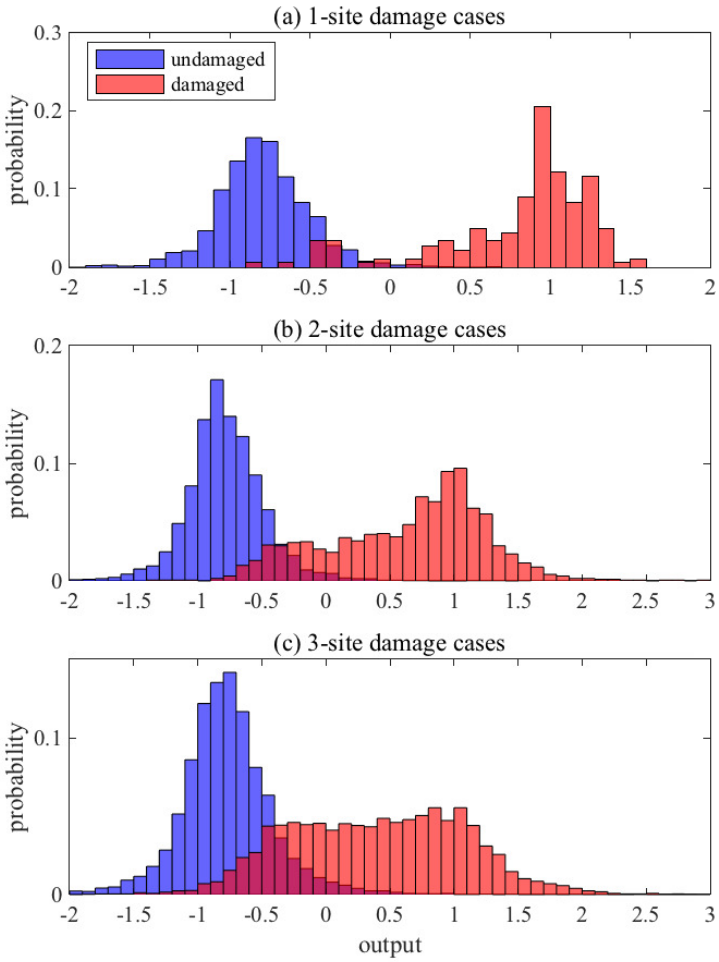


Figure 4.13. Combined histograms of MLC outputs. y-label is set as "probability" instead of "frequency" for better presentation.

Figures 4.13 (a) to (c) illustrate the distribution of "undamaged" and "damaged" labels for the cases with NDS = 1 to 3. The bin sizes are set identical for better compar-

ison. It can be observed that MLC has negative outputs for most "undamaged" labels and some of the "damaged" labels. This indicates that, similar to BC, MLC has better labeling performance for "undamaged" cases than for "damaged" cases, though it can provide better overall classification result as indicated from the scores in Table 4.4. However, thresholds for decision-making should be adjusted in favor of true-positive rate by sacrificing the true-negative rate for some special applications such as clinical diagnostic tests [110]. Adjusting decision thresholds to account for varied misclassification costs and/or prior probabilities has been studied via ROC (receiver operating characteristic) analysis [111, 112, 113]. Health monitoring of important infrastructures needs relatively conservative recommendations about the structural damage states. Therefore, the threshold used for classification can be slightly reduced to increase the true-positive rate while reducing the true-negative rate for damage identification. For example, reducing the threshold from the default value 0 to a negative value, say -0.5, can improve the classification accuracy of "damaged" labels at the sacrifice of the accuracy of "undamaged" labels.

Figures 4.14 (a) to (f) show the same data as in Figure 4.13 with variable bin sizes and three different threshold values. It shows that, reducing the threshold can significantly increase the accuracy of "damaged" labels and keep the accuracy of "undamaged" labels generally acceptable. For example, when the threshold is reduced from 0 to -0.5, the accuracy of the "damaged" is increase from 0.68 to 0.91 (an increase of 0.23) for NDS = 3, and the accuracy of the "undamaged" still keeps above 0.8 though reduced by 0.15. Hence, the threshold adjustment can largely improve the general performance of MLC for multi-site structural damage identification, which is not plausible in MCC. The final threshold can be determined by balancing the benefits of the increased true-positive rate and the loss

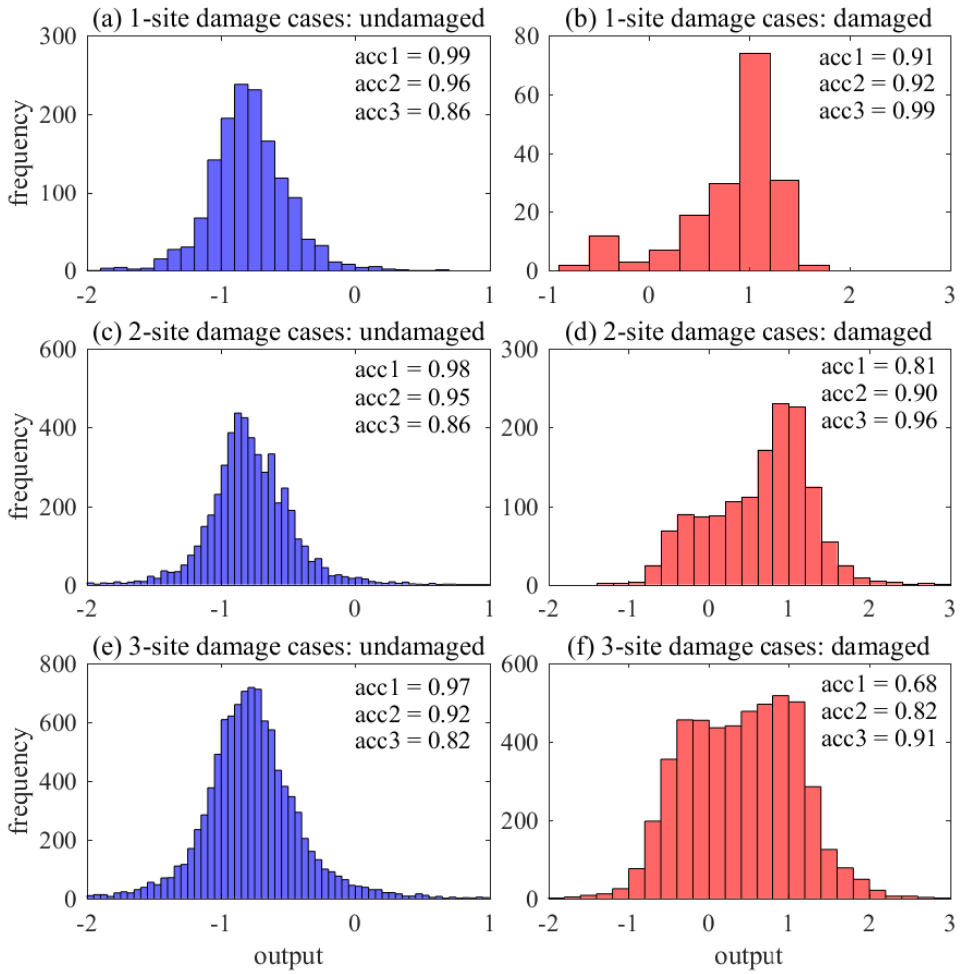


Figure 4.14. Separate histograms of MLC outputs. The accuracies are obtained with different thresholds: acc1 for a threshold of 0; acc2 for -0.3; acc3 for -0.5.

caused by the reduced true-negative rate.

Figures 4.15 (a) and (b) present the histograms of MLC outputs when testing the trained classifier on 3-site damage cases that are not included in the training dataset. One can find that the testing accuracy on unobserved "damaged" labels is lower than that on observed "damaged" labels as shown in Figure 4.14 (f). When the decision threshold is adjusted to -0.5, both "damaged" and "undamaged" labels have an accuracy above or equal to 0.8. It is noted that satisfactory damage identification of unobserved damage cases is a significant advantage of MLC over MCC and BC.

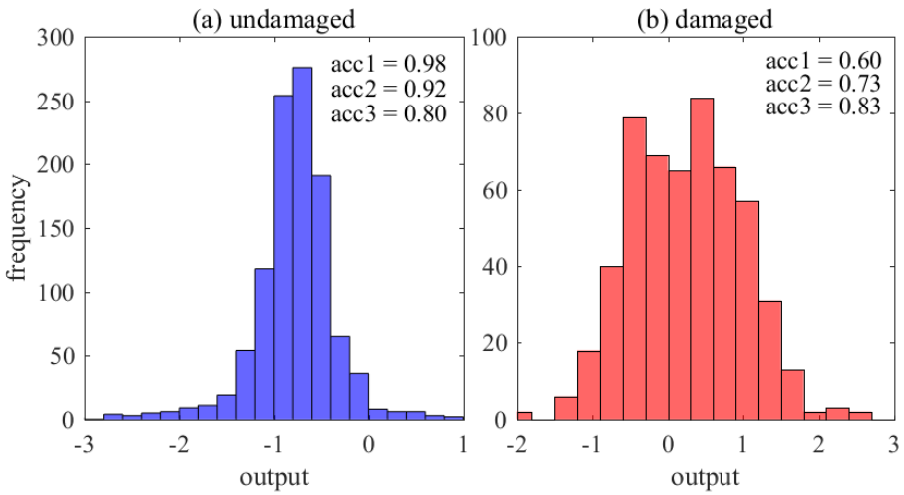


Figure 4.15. Histograms of MLC outputs when tested on unobserved 3-site damage cases.

4.5. Experimental Validation

This section uses experimental data to examine the effectiveness of the MLC method in identifying multi-site structural damage and its advantage over the MCC and BC methods. This study uses the data published by Los Alamos National Laboratory for a 4-DOF frame structure [23] as depicted in Figure 4.16. Damage was introduced to the frame structure by stiffness reduction of columns and mass addition of floors to simulate the operational and environmental influence, and bumper impacts to simulate the nonlinear behaviors of fatigue cracks. An electrodynamic shaker excited the frame structure with various damage conditions using the Gaussian white noise on the base floor along the structure's centerline. The excitation force from the shaker to the structure was measured with a load cell mounted on the stringer and the structural acceleration response was measured using four accelerometers attached at the edge center of each floor. The data acquisition sampling frequency was 320 Hz. For each damage state, ten shaking tests were conducted considering the variability of excitations and structural prop-

erties. The recorded data of excitations and structural responses can be used to extract damage-sensitive features for structural health monitoring purposes [23].

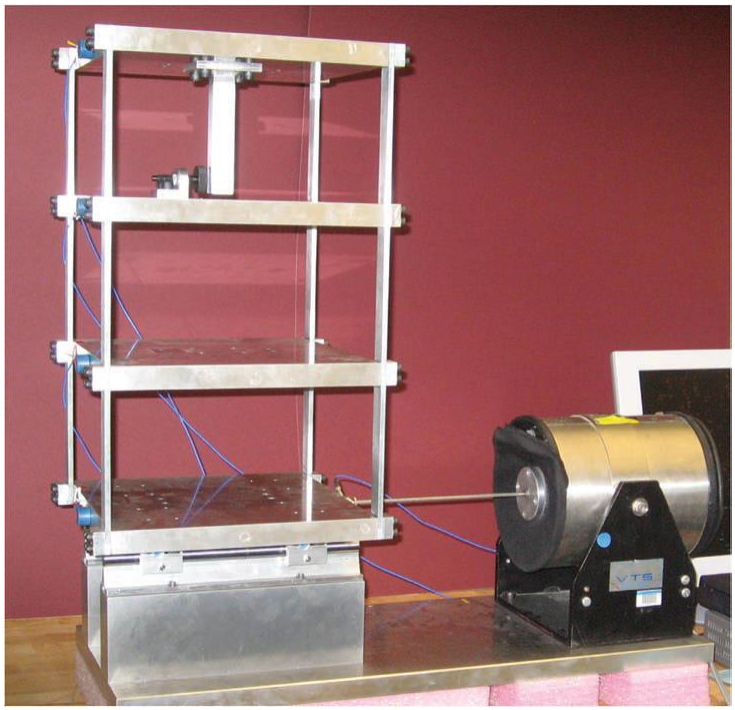


Figure 4.16. Three-story building structure and shaker [23]. Damage are introduced by either adding mass on the first floor and/or introducing gap to the bumper on the third floor.

Table 4.5. MLC scores from *InsDif*-MLC and that integrated from BC outputs.

evaluation metric	MLC	BC
hamming loss ↓	0.14	0.25
one error ↓	0.06	0.22
coverage ↓	0.34	1.31
ranking loss ↓	0.06	0.34
average precision ↑	0.97	0.85

The present study selects three damage cases from the database available in [23] that fit in the scenario of multi-site damage detection in this paper: (1) mass addition; (2) gap introduction; (3) mass addition and gap introduction. The first two cases, i.e., mass addition and gap introduction, respectively, represent two single-site damage cases, and

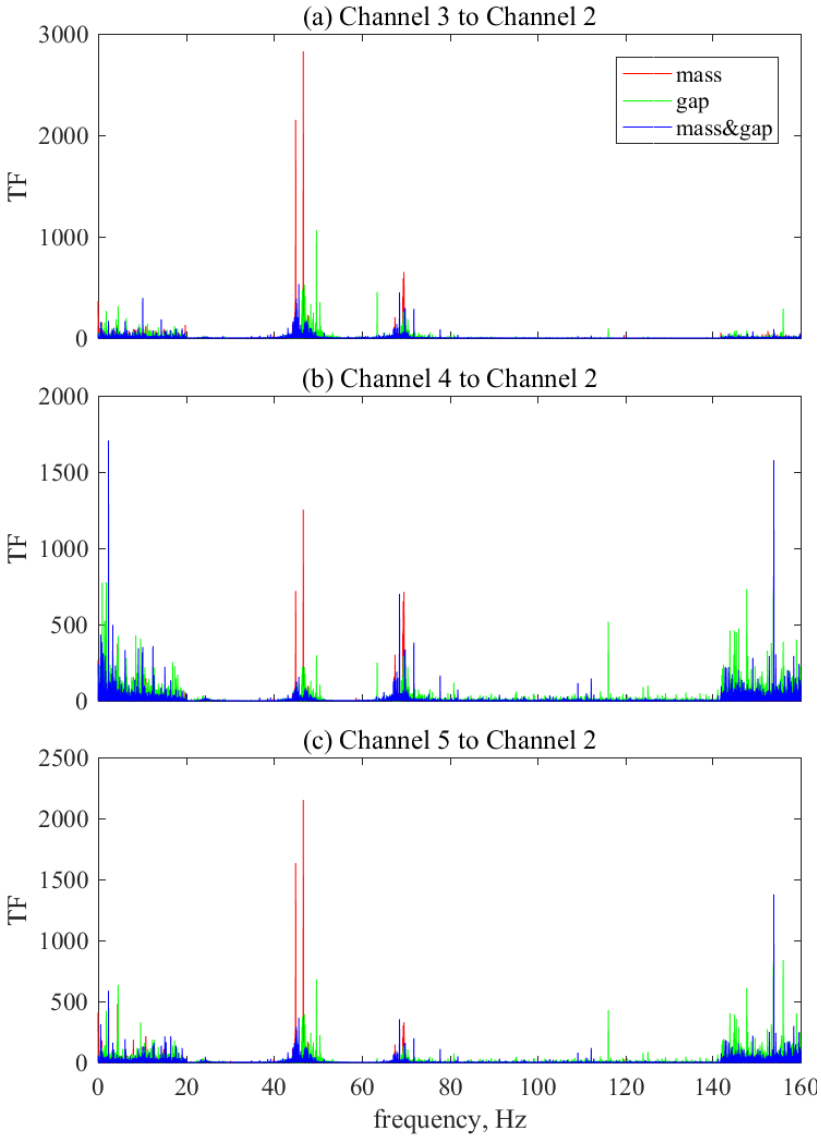


Figure 4.17. Transmissibility functions of 3 damage states

the third case, i.e., mass addition plus gap introduction, stands for the case with multi-site damage. In cases with mass addition, a 1.2 kg mass is added to the first floor, which equals 19% of the original mass of the first floor (6.279 kg); In cases with introduced gap, the gap size ranges between 0.05mm to 0.20mm. More details about the experimental setup can be found in [23]. This section uses MCC, BC, and MLC methods to identify damage conditions and evaluates their performance.

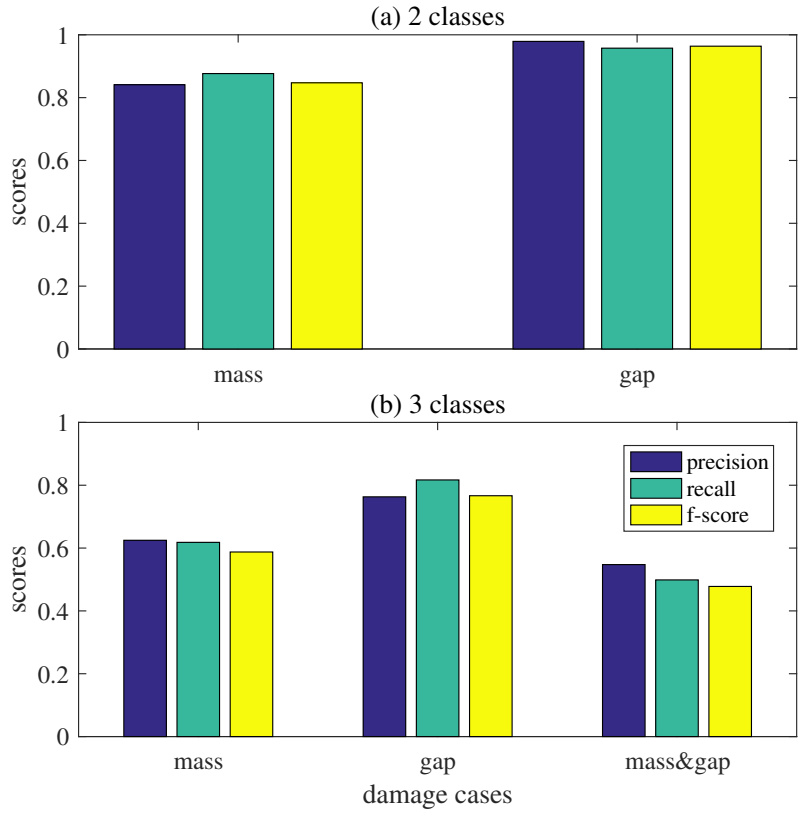


Figure 4.18. MCC scores for the two scenarios.

This experimental study uses transmissibility functions (TFs) as features for damage identification. TFs are calculated via the point-wise ratio of the response spectrum between two different channels measuring the responses of two DOFs. Figures 4.17 (a) to (c) show the TFs of channels 3, 4, and 5 to channel 2. Before used for training, the TF features are reduced in dimension using the PCA technique to avoid the curse of dimensionality and improve the trained models' generality.

This section considers two scenarios for MCC analysis. The first scenario contains two classes that correspond to the two single-site damage cases, respectively. This scenario is designed to show the MCC performance on two exclusive cases without the ambiguity introduced by the multi-site damage case that shares one damage site with each of them.

In comparison, the second scenario contains all the three cases to demonstrate the potential influence of ambiguity on the classification quality. Figures 4.18 (a) and (b) show the classification scores of each class in the two scenarios. Figure 4.18 (a) shows that without the interference of case (3), MCC yields high scores for both case (1) and case (2) with all the three scores above 0.8. However, the ambiguity brought by case (3) reduces the classification quality of both case (1) and case (2) and leads to deficient classification performance on case (3) itself. Hence, MCC ignores the correlation between the single-site and multi-site damage cases and thus fails to resolve the ambiguity between them, which can largely impair the classification performance on all cases.

It is noted that each 1-site damage case has one positive label "+1" for the damaged site and one negative label "-1" for the undamaged. In comparison, the two-site damage case has a positive label for both the two sites. Table 4.5 compares the scores from *InsDif*-MLC and that calculated from BC outputs. It shows that MLC outperforms BC with better scores for all the evaluation metrics. Figures 4.19 (a) and (b) present the histograms of MLC outputs labeled with the corresponding accuracies. One can see that the 1-site damage case has a higher accuracy for "damaged" case than for "undamaged" case because more positive labels are included in the training data from the experiment. It is shown that the accuracies for all conditions are above 0.8 without adjusting the decision threshold. Hence, using experimental data, this section demonstrates that the ambiguity in multi-site damage cases can reduce the damage identification accuracy without advanced data processing. It also highlights the advantage of the MLC method over MCC and BC methods in identifying multi-site damage. Moreover, labeling accuracy is largely dependent on the availability of labels in the training dataset.

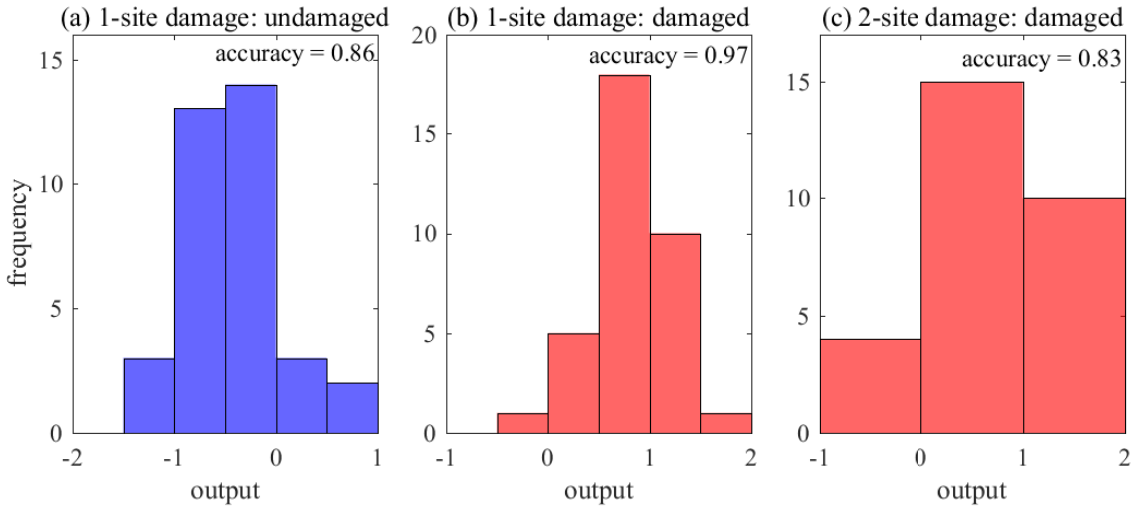


Figure 4.19. Histograms of MLC outputs.

4.6. Summary

Multi-site damage identification has significant importance in structural health monitoring. The main challenge faced by multi-site damage detection lies in the ambiguity caused by information mixture when damage happens at multiple locations. To address this challenge, the present study proposes a data-driven method using a multi-label classification (MLC) method and the instance differentiation algorithm (*InsDif*). MLC denotes a damage case having multiple damage sites with an instance containing multiple labels. The *InsDif* algorithm transforms the multi-site damage identification issue from a single-instance multi-label problem to a multi-instance multi-label problem to discriminate the labels containing mixed information. The algorithm explicitly expresses the ambiguity in the input space by transforming the examples into a bag of instances with each corresponding to one class label. A k-medoid clustering helps discover the underlying information hidden in a dataset with superficial obscurity. Decomposition of the inter-correlation between instances minimizes the confusion when training the classifier, which makes MLC

superior to MCC and BC in detecting multi-site damage.

The presented numerical and experimental results approves the impaired performance of damage identification resulting from the ambiguity among cases. Multiclass classification (MCC) has limited generality on the testing examples especially when the data quantity is not large enough, as it regards all the damage cases, though sharing damage sites, as exclusive categories and thus ignores the correlation among them. Additionally, MCC cannot label correctly an example with unseen labels in the model training and does not support label-by-label evaluation, i.e., site-by-site damage detection. Binary classification (BC) is a reduced scheme of MLC that trains a separate classifier for each possible damage location and finally integrates the outputs of all classifiers for an overall structural damage evaluation. Though it evaluates each site separately, BC fails to account for the correlation among cases sharing damage sites and thus has limited performance in identifying the damage occurrence. In comparison with BC and MCC, MLC attempts to resolve the ambiguity and thus yields more encouraging results. Scatter-plots of the raw features and the transformed features by the *InsDif* algorithm approves the advantage of MLC in disentangling the mixed information contained in the input space. Numerical analysis shows that the damage identification accuracy can be largely improved by adjusting the decision threshold when determining the category of testing examples. The result of experimental analysis validates the effectiveness MLC and indicates that MLC outperforms MCC and BC in identifying multi-site damage occurrence. Moreover, MLC preserves generality when tested on unobserved multi-site damage cases, which indicates that it has the potential of mitigating the critical data insufficiency problem in data-based structural health monitoring.

Chapter 5. A Numerical Study on Multi-Site Damage Identification: A Data-Driven Method via Constrained Independent Component Analysis

5.1. Introduction

In Chapter 4, the multi-label classification method is investigated to improve the performance of multi-site damage detection, in which the *InsDif* algorithm is used to decode the mixed information contained in the damage cases. To further relieve the burden from data insufficiency, the present chapter investigates the constrained independent component analysis and its application in data-driven multi-site damage detection.

In the community of structural health monitoring (SHM), multi-site damage identification is important and challenging while receiving inadequate research effort especially in the area of data-based SHM [5]. Engineering structures are inevitable to experience multi-site damage during their service life. Therefore, it is essential to develop efficient methodologies for multi-site damage identification.

In model-based methods, multi-site damage identification does not involve large extra efforts compared to single-site damage identification. Model-based methods identify structural damage via using model updating techniques, which optimize the structural parameters that represent the damage locations and severities to minimize the objective function. The model-based method necessitates the implementation of model-updating each time when damage identification is needed, which is always computationally expensive and sometimes ill-posed. Challenges of this method also come from uncertainties of physical variability, experimental validity, and model-form error [5]. Contursi et al. optimized the multiple damage location assurance criterion (MDLAC) using the traversal

search strategy and successfully located the multiple structural damage [81]. Sohn and Law identified multi-damage occurrence via using a Bayesian-based branch-and-bound search scheme to optimize an objective function based on certain modal properties [47]. Li et al. numerically and experimentally investigated using modal-energy-based method to localize multi-site damage on wind turbine structures [114]. Hosseinzadeh et al. detected and evaluated multi-site structural damage through optimizing an objective function based on modal assurance criterion (MAC) and modal flexibility matrix using the democratic particle swarm optimization (DPSO) algorithm [115].

In comparison, the data-based method is based on a statistical rather than a physical model of the structure. It generalizes damage identification to a supervised/unsupervised learning problem in the statistical pattern recognition paradigm. Compared with model based methods, this method possesses the advantage in computational efficiency and uncertainty handling [5, 116]. The major drawback of the data-based SHM is its requirement on structural data for all possible damage states, which is rarely available in practice. It becomes even more critical when considering statistical pattern recognition for multi-site damage identification. A straightforward data-driven approach requires collecting structural damage data for all possible combinations of damage locations. Hence, the number of damage cases covering all combinations would grow exponentially with the number of possible damage locations. Therefore, it is apparent that the data insufficiency problem can be a critical issue in data-driven multiple-site damage location. There are limited studies on multi-site damage identification using data-driven methods. Barthorpe and Worden [5] used a data-based approach to identify multi-site damage on an aircraft wing with data only from single-site damage. This study general-

izes the damage detection and localization problems to a machine learning classification paradigm and trains distinct classifiers for identifying the damage condition at each location of interest. In addition, Zhang and Sun [117] attemped to improve the performance of multi-site structural damage identification in data-driven SHM via multi-label classification (MLC). It shows that compared with other machine learning schemes including multi-class classification (MCC) and binary classification (BC), MLC yields much higher accuracy in multi-site damage identification, as it accounts for the physical correlation between damage cases sharing damage locations and thus reduces the confusion in the learning process.

As an effective technique, blind source separation has been widely used for system identification [118, 119, 120, 121, 122] and has the potential for structural health monitoring. Independent component analysis (ICA) is probably the most widely used method of performing blind source separation [123]. Assuming that most measured signals $\mathbf{x}(t)$ are a mixture of independent signals $\mathbf{s}(t)$, ICA finds a linear transformation yielding signals that are as independent from each other as possible from a statistic perspective. That is $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) = \sum_{i=1}^{n} \mathbf{a}_i s_i(t)$ [124], where $\mathbf{A}$ is the mixing matrix and $\mathbf{a}_i$ are its columns. The ICA-based method of blind source separation estimates $\mathbf{s}(t)$ and $\mathbf{A}$ using the maximum likelihood method based on the classical central limit theorem. The best estimation of $\mathbf{s}(t)$ has the largest possible statistical independence and non-Gaussianity, which can be measured by kurtosis or negentropy [123].

As ICA has been proved efficient in analyzing structural dynamic responses [125, 92, 126], it should have the potential for vibration-based structural health monitoring by aiding the extraction of damage sensitive features. Given structural responses, ICA yields

the mixing matrix $\mathbf{A}$ and the independent components (ICs) $\mathbf{s}(t)$, which contains the information of structural damage and thus can be used for extracting damage-sensitive features.

The ICs have been used as features in data-driven structural damage identification which yields high accuracy [127]. However, from the perspective of pattern recognition, the dimension of the IC-based damage features are inappropriately large and thus may cause overfitting issues due to the curse of dimensionality [12]. Additionally, taking the ICs as the single source of damage features may reduce the damage detection quality because the potential damage information contained in the mixing matrix is discarded without rigorous verification. In comparison with ICs, the mixing matrix has compact dimension, which can largely alleviate the cumbersomeness of feature extraction and circumvents the overfitting issue caused by feature dimensionality [128]. Ref. [124] uses both the ICs and the mixing matrix for online multi-site damage detection and localization.

In summary, ICA has the potential for data-driven structural damage identification due to its efficiency in analyzing structural dynamic responses. Either its resulting ICs or mixing matrix can be directly used as damage features or further analyzed for extracting advanced damage-sensitive features. However, traditional ICA algorithms do not guarantee identical output ICs due to the issue of local minimum [129, 130, 120, 131], which means that the ICs and mixing matrices are randomly permuted or shifted for each run. Therefore, it becomes challenging to extract damage-sensitive feature that captures the most valuable information using either the ICs or the mixing matrix.

On the other hand, the ICA outputs can be regulated or directed by using the cICA. cICA [132] or ICA with reference (ICA-R) [133] imposes constraints on the output

of ICA decomposition referring to the reference signals that are established with *a priori* knowledge. The pre-specified reference signals that represent a rough template of the desired ICs direct the decomposition algorithm towards matching the references. cICA avoids estimating the number of sources and reduces the computational cost in case there are a large number of sources [134]. cICA has applications in mechanical fault diagnosis with reference signals constructed by virtue of the prior knowledge of machine faults, and it has been proved to be advantageous over traditional ICA methods [134, 135].

This study aims to develop an efficient methodology for single- and especially multi-site damage identification via using data-driven methods and cICA algorithms. As aforementioned, it is important to reduce the feature dimension yet it is also challenging to preserve all the valuable information related to structural health condition when conducting dimension reduction. To address this challenge, the present paper proposes to extract damage-sensitive features from the cICA outcome on recorded structural responses. The novelty of the present study is twofold. First, a new methodology is proposed to identify single-/multi-site damage using damage-sensitive features derived from cICA. Second, appropriate damage-sensitive features have been selected, which can provide accurate and efficient damage identification results. By enforcing the constraint of identical output ICs, the damage-related information contained in the dynamic responses are compacted into the low-dimension mixing matrices. Subsequently, feature extraction from the mixing matrix can then be less complex. A case study is carried out on an 8-DOF lumped mass system. Stiffness reduction was introduced to 1-3 springs to simulate single-site and multi-site damage. Harmonic excitations are applied in the present study. However, it is noted that the proposed methodology is not limited to harmonic excitations. Impact

loading also works for this purpose, and the results will be presented in future study for experimental validation. In addition, some service/ambient loading (e.g., traffic loading on bridges) composed of harmonics with multiple frequencies has the potential for multi-site damage identification using the proposed methodology. The research results show that the features extracted from cICA can help identify the single-site damage and multi-site damage progressively. More importantly, the proposed method has the potential to locate multi-site damage using single-site damage data only, excluding the prerequisite of having multi-site damage data as references. This merits largely relieves the data insufficiency issue in data-driven structural health monitoring.

5.2. Problem Statement and Methodology

5.2.1. Problem Statement

Identifying structural damage locations requires features that are sensitive to the damage location and independent of its severity. While proved efficient in localizing single-site damage, many features may not work for multi-site damage identification. The frequency change ratio has been proved independent of the damage severity but its occurring location [90], and thus it has been widely used for data-driven structural damage identification [25]. When stiffness reduction happens at an unknown location m on the structure, the frequency changes of the i^{th} and j^{th} modes are:

$$\Delta f_i = \frac{\partial f_i}{\partial k_m} \Delta k_m \tag{5.1}$$

and

$$\Delta f_j = \frac{\partial f_j}{\partial k_m} \Delta k_m \tag{5.2}$$

where f_i and f_j are the natural frequencies of the i^{th} and j^{th} modes; k_m denotes the structural stiffness at location m; $\frac{\partial f_i}{\partial k_m}$ is the sensitivity of f_i to k_m variation. It is noted that $\frac{\partial f_i}{\partial k_m}$ is a function of the damage location (i.e., k) and independent of its severity Δk_m. Taking the ratio between Δf_i and Δf_j yields:

$$\frac{\Delta f_i}{\Delta f_j} = \frac{\frac{\partial f_i}{\partial k_m}}{\frac{\partial f_j}{\partial k_m}} \tag{5.3}$$

which cancels out the effect of damage severity Δk_m and preserves the information about the damage location.

However, for a multi-site damage case, the equations become:

$$\Delta f_i = \frac{\partial f_i}{\partial k_1}\Delta k_1 + \frac{\partial f_i}{\partial k_2}\Delta k_2 + ... + \frac{\partial f_i}{\partial k_N}\Delta k_N, \tag{5.4}$$

$$\Delta f_j = \frac{\partial f_j}{\partial k_1}\Delta k_1 + \frac{\partial f_j}{\partial k_2}\Delta k_2 + ... + \frac{\partial f_j}{\partial k_N}\Delta k_N, \tag{5.5}$$

and

$$\frac{\Delta f_i}{\Delta f_j} = \frac{\frac{\partial f_i}{\partial k_1}\Delta k_1 + \frac{\partial f_i}{\partial k_2}\Delta k_2 + ... + \frac{\partial f_i}{\partial k_N}\Delta k_N}{\frac{\partial f_j}{\partial k_1}\Delta k_1 + \frac{\partial f_j}{\partial k_2}\Delta k_2 + ... + \frac{\partial f_j}{\partial k_N}\Delta k_N} \tag{5.6}$$

One can find that the ratio of frequency change cannot exclude the influence from the damage severities $\Delta k_m (m = 1, 2, ...N)$, unless they are mutually proportional, which is hardly possible in real situation. This fact invalidates the features based on frequency changes for multi-site damage identification. Similar issues can happen to many other modal-based features [91, 25]. Therefore, the identification of multi-site damage necessitates extracting appropriate features that can reflect the contribution from each potential damage location to the resulting structural responses.

5.2.2. Methodology

In data-driven methods for SHM, feature extraction aims to produce damage-sensitive features with reduced dimensions. As an efficient technique, ICA can resolve the structural responses into a specified number of ICs that are linearly combined by the mixing matrix. As a result, the structural damage information contained in the responses are transferred to the the mixing matrix and the ICs. Compared with the ICs, the mixing matrix has compact dimensions and thus is more attractive for feature extraction.

Referring to Refs [129] and [132], the present study proposes using cICA to extract

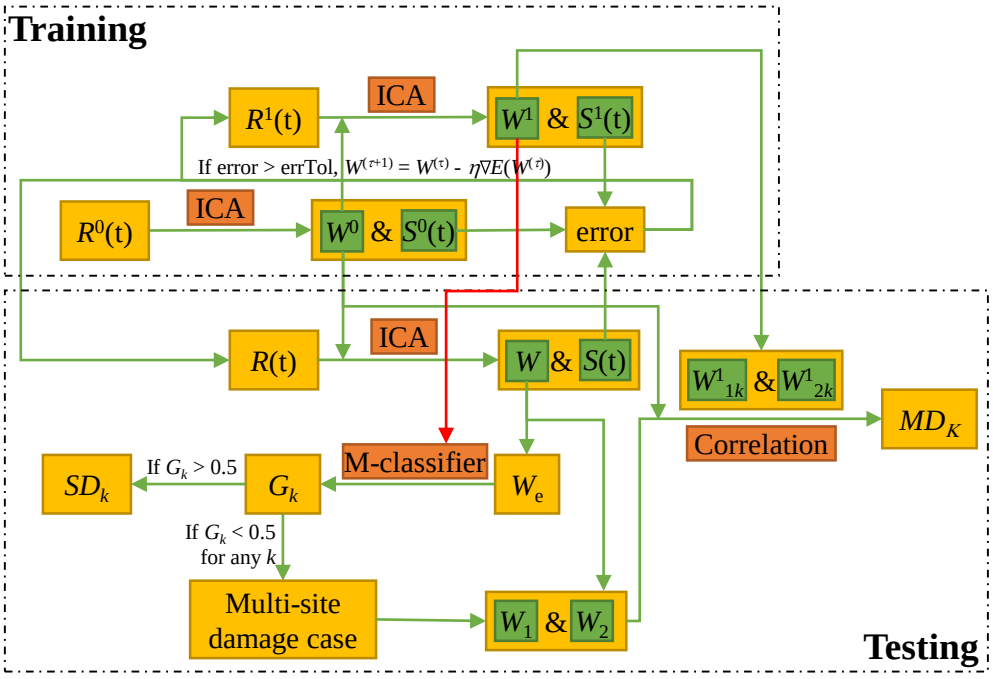


Figure 5.1. Flowchart of damage identification using constrained ICA (cICA). $R^0(t)$: responses from an intact structure; $R^1(t)$: responses from structures with single-site damage; $R(t)$: responses from a structure with an unknown damage state. ICA: independent component analysis. W^0: mixing matrix from $R^0(t)$; W^1: mixing matrix from $R^1(t)$; W: mixing matrix from $R(t)$. $S^0(t)$: independent components (ICs) from $R^0(t)$; $S^1(t)$: ICs from $R^1(t)$; $S(t)$: ICs from $R(t)$. errorTol: tolerance of error for convergence. η: learning rate. $\nabla E(W^{(\tau)})$: gradient vector at step τ. W_e: elements of W. G_k: classification score. SD_k: single-damage case at location k. W_1 and W_2: columns of W. W^1_{1k} and W^1_{2k}: columns of W^1 with single-site damage at k. MD_K: multi-damage cases at locations in vector K.

compact and informative features from structural responses for single- and multi-site damage identification. Harmonic loading is used so that the responses under different damage states are comparable. Under the simultaneous excitation of two harmonic loading with different frequencies, the structure will experience steady-state responses with the two frequency components. Subsequently, ICA on the responses will yield two harmonic ICs corresponding to the two harmonic loading. If the output ICs are regulated to be identical, the mixing matrix will contain more structural information than that obtained using one harmonic loading as in [129]. It should be noted that this methodology is not limited to harmonic loading. Any loading groups that generate indepdent dynamic responses should work. This analysis together with the experimental study will be left for future work due to the length limitation of this paper. Fig. 5.1 depicts the key procedures of the proposed methodology. The FastICA algorithm [136] is used in this study for all the ICA implementations.

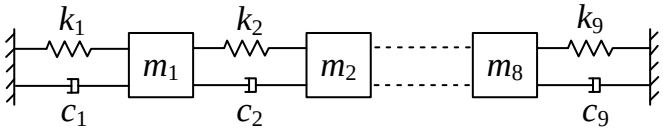


Figure 5.2. 8-DOF lumped mass system for numerical simulation. $m_i(i = 1, 2, ..8)$: masses; $c_i(i = 1, 2, ..9)$: dampers; $k_i(i = 1, 2, ..9)$: springs.

As shown in Fig. 5.1, the proposed methodology consists of two modules: a training module and a testing module. The training module uses the responses of an intact structure referred to as $R^0(t)$ and that of the structures with single damage at specified locations referred to as $R^1(t)$. To improve the generality and robustness of the proposed methodology, all the responses in this study are polluted by a white Gaussian noise that has a peak of 5% of the response. The first step is to conduct ICA on $R^0(t)$ with the num-

ber of ICs set as two. Due to the instability of the FastICA algorithm and the noise interference, identical output ICs cannot be guaranteed. Therefore, the initial step needs to be repeatedly run until the resulting ICs become two harmonic series with different frequencies. Then the output mixing matrix and the ICs are recored as W^0 and $S^0(t)$, respectively.

Table 5.1. Mechanical properties of the 8-DOF model

ID	1	2	3	4	5	6	7	8	9
mass	1	2	3	4	2	2	1	3	-
damping coefficient	20	20	20	20	20	20	20	20	20
spring stiffness	10^4	10^4	10^4	10^4	10^4	10^4	10^4	10^4	10^4

The second step is ICA on the responses of single-damage structures $R^1(t)$. To ensure identical ICs with that of $R^0(t)$ within the allowed tolerance (i.e., errTol in Fig. 5.1), the mixing matrix is updated using the gradient optimization scheme in each iteration [137]:

$$W^{(\tau+1)} = W^{(\tau)} - \eta \nabla E(W^{(\tau)}) \tag{5.7}$$

where τ denotes the iteration step; η denotes the specified learning rate; $\nabla E(W^{(\tau)})$ denotes the gradient vector at step τ. This iteration converges fast owing to the adoption of W^0 as the reference mixing matrix. Once the ICs of $R^1(t)$, $S^1(t)$, converges to $S^0(t)$, the corresponding mixing matrix is termed as W^1. It is noted that the cICA on $R^1(t)$ needs to be implemented on all the single-damage cases that are of monitoring interest to provide sufficient reference data to train the identification model.

The third step trains a classifier based on the Mahalanobis distance, M-classifier, using the single-site damage cases in the training module with the exclusive damage loca-

tions set as target class labels. As shown in Fig 5.1, the cICA on $R^0(t)$ and $R^1(t)$ and the classifier training constitute the training module of this methodology.

Table 5.2. Properties of harmonic excitations

Load case	Load ID	Loading location	Magnitude	Circular frequency
I	1	mass 2	10^6	1
	2	mass 7	10^6	2
II	1	mass 3	10^6	1
	2	mass 6	10^6	2

In the testing module, $R(t)$ is the response of the structure with an unknown damage condition under the same harmonic excitation as used for generating $R^0(t)$ and $R^1(t)$ in the training module. cICA is performed on $R(t)$ by setting W^0 as the reference mixing matrix and updating it for identical output ICs. The elements of the resulting mixing matrix, W_{e}, are entered into the trained M-classifier, which yields a classification score for each target class based on its distance to the reference data therein. If the classification score G_k of class k is larger than 0.5 for a certain class k, i.e., $G_k > 0.5$, this damage state is evaluated as a single-damage case at location k; otherwise, it is determined as a multi-site damage case, and the specific damage locations will be further analyzed. It should be noted that the classifier threshold 0.5 is a generic value. The specific threshold of the M-classifier needs to be determined based the distribution of Mahalanobis distances between damage cases.

If determined as a multi-site damage case, the two columns of the mixing matrix, referred to as W_1 and W_2, corresponding to the two ICs respectively will be used for further analysis. The correlation of W_1 and W_2 with that from a reference single-site damage case (W^1_{1k} and W^1_{2k}) is calculated. Numerical results show that the correlation coefficients

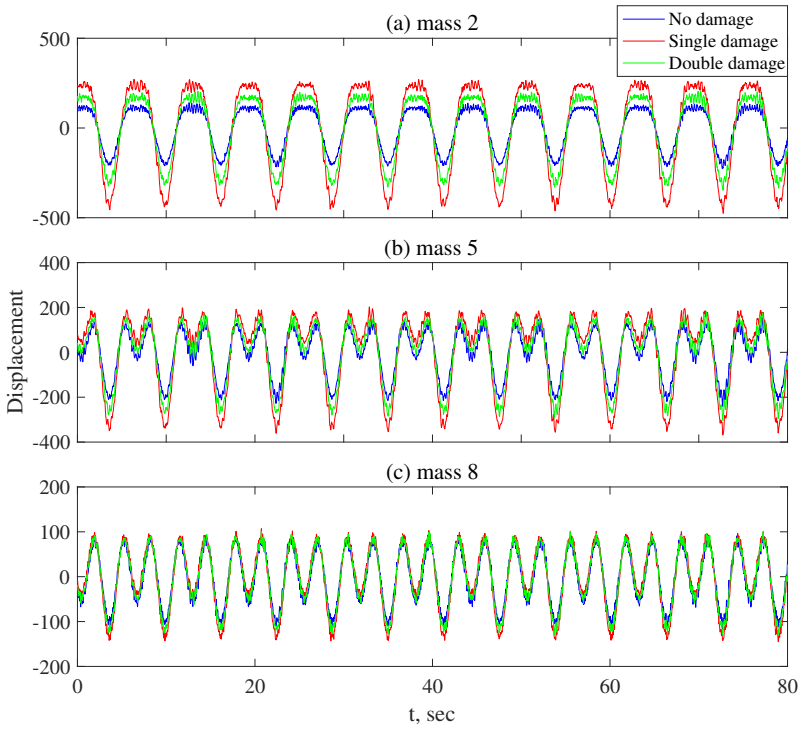


Figure 5.3. Comparison of structural responses with different damage states. (a) the displacement of m_2; (b) the displacement of m_5; (c) the displacement of m_8.

will reach the local maximum when k is one of the multi-site damage locations. Hence, the multi-site damage can be identified via comparing the correlation coefficients of W_1 and W_2 with W_{1k}^1 and W_{2k}^1, respectively. The merits of this method include: first, it identifies the structural damage location(s) progressively by evaluating whether it belongs to a single-damage case or a multi-damage case and then determining the damage locations; second, this method accomplishes multi-site damage detection with reference to merely single-site damage cases, obviating the prerequisite of having multi-site damage data in the training module, which is always a challenge of data-based SHM. The next section will present a detailed numerical study to demonstrate and verify the proposed method.

5.3. A Numerical Case Study

This section presents a numerical case study to examine the effectiveness of the proposed methodology for single- and multi-site structural damage identification.

5.3.1. Numerical Model

The present study uses an 8-DOF lumped mass system as shown in Figure 5.2. Table 5.1 lists the non-dimensional parameter values of the mechanical properties. This case study simulates the structural damage via a fractional loss of the spring stiffness, which are randomly specified from 0.1 to 0.9 to account for variations in structural damage severity. Dynamic tests in this study apply two harmonic loading of disparate frequencies to two masses of the lumped system and collect the displacement responses of all the masses. Table 5.2 lists the information of the two load groups used in this study. A preliminary sensitivity study indicates that a minimum of two responses excited by the harmonic loading are required for the ICA implementation to ensure satisfactory damage identification performance. In addition to the randomization of the damage severities, this study adds white noise of 5% in magnitude to the collected responses to resemble the uncertainties from environmental effects and signal collection. Fig. 5.3 shows the displacement of masses 2, 5, and 8 under an intact condition and damage conditions with single and double damages. One can find that each response curve contains multiple frequency components corresponding to the two excitation frequencies and the random noise.

5.3.2. Single-Site Damage Identification

This subsection focuses on single damage identification. As illustrated in Fig. 5.1, the responses from the intact structure, $R^0(t)$, are processed first using the FastICA algo-

rithm. The number of target ICs is set as two in order to extract the two frequency components contained in the response series. Fig. 5.4 shows the output ICs from three separate runs of FastICA with the same settings. It can be observed that the first run yields a harmonic IC (IC 1) and an irregular series (IC 2), and both the second and third runs produce two harmonic ICs with different frequencies. Actually, the harmonic ICs from the second and third runs are identical but numbered reversely, which will cause column exchange in the corresponding mixing matrices and thus inconvenience for comparison. This result indicates that the ICA on a given input will generate disparate ICs and mixing matrices during different runs. As FastICA seeks the signal components having the least Gaussianity by optimizing the transformation matrix via Newton iteration [123], it may sometimes fall into local minima especially in cases with random disturbance. Therefore, when used for structural health monitoring, proper constraint should be imposed on ICA so that their outputs are comparable for extracting damage-sensitive features. In this study, the ICA on $R^0(t)$ is iterated until it yields two harmonic components with different

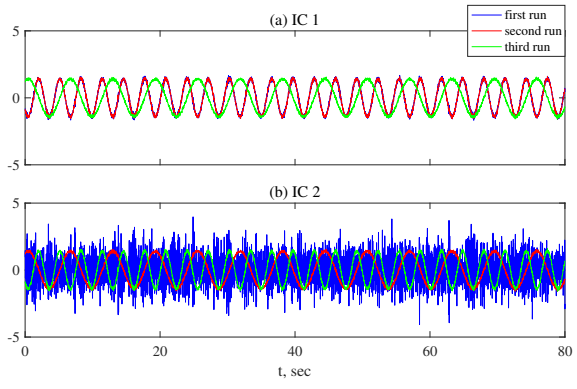


Figure 5.4. Independent components (ICs) from three separate simulations. (a) the first independent component IC 1. (b) the second independent component IC 2.

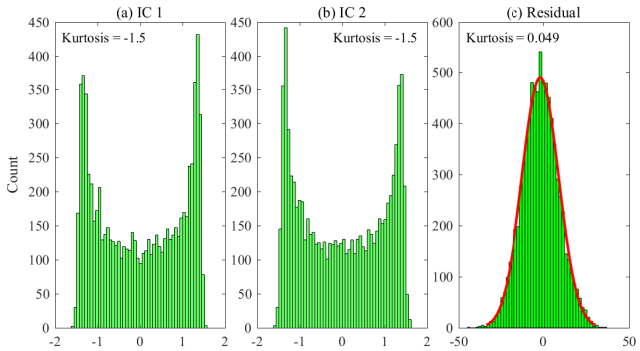


Figure 5.5. Statistical distribution of independent components (ICs) and residual using the responses from the intact structure. The red curve in (c) is the fitted normal distribution.

frequencies as expected. The quality of ICA can be evaluated by the Gaussianity of the output ICs and its residual. Fig. 5.5 compares the distributions of IC1 and IC2 from $R^0(t)$ and the residual. One can find that both ICs follow a non-Gaussian distribution. Compared with the ICs, the residual follows a Gaussian distribution as shown in Fig. 5.5(c). This verifies that FastICA can separate the two most non-Gaussian components from the Gaussian residual.

Another quality of the ICs is their capability of representing the input signals. For structural damage identification, the ICs that are most representative of the response characteristics are preferred, especially when the number of ICs is set smaller than that of the response series. Fig. 5.6 compares the reconstructed responses from the output ICs with the original simulated responses of the intact structure. One can observe that the reconstructed responses match well with the original responses, which means that the harmonic ICs captures the foremost features of the structural responses, leaving the Gaussian component in the residual with the least valuable information.

Following the flowchart in Fig. 5.1, this case study subsequently introduces damage to the structure and conducts dynamic tests using identical harmonic excitations as used on the intact structure. Then the collected responses are processed using the constrained ICA (cICA) to generate approximately identical ICs and distinct mixing matrices that contain the information about the structural damage states. Fig. 5.7 shows that cICA on the responses from damaged structures yields ICs that are consistent with the harmonic ICs from the intact structure. Considering the identical excitations used in the dynamic tests, it is expected that the valuable information regarding the structural changes reflected by the responses are compacted into the mixing matrices via using cICA.

Before training a machine learning model using the extracted features, it is recommended to evaluate their capability of distinguishing instances belonging to different target classes. The mixing matrix in this case study has a dimension of 8×2, corresponding to the eight DOFs and the two ICs. For the sake of visualization, the feature vectors are reduced to two principal components (PCs) using the principal component analysis (PCA). PCA preserves the feature components with the largest variances [138, 139, 140]. Figs. 5.8 (a) to (c) illustrate the two principal components (PC 1 and PC2) corresponding to the nine single-site damage cases referred to as SD1-SD9 and that with the double-site

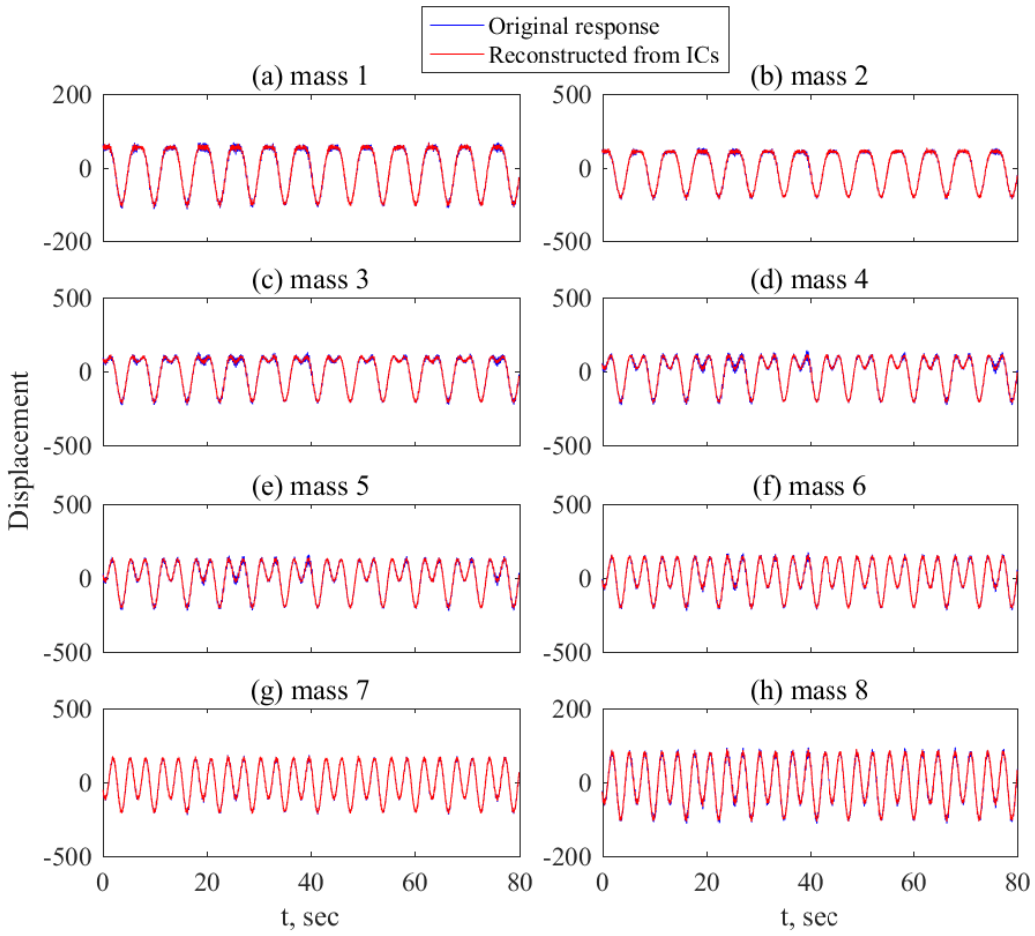


Figure 5.6. Comparison of measured and reconstructed structural responses. (a) to (h): the displacement time-history of masses 1 to 8.

damage cases. Fig. 5.8 (a) shows that PC1 and PC2 corresponding to the nine single-site damage cases display explicit distinction. However, the double-site damage cases cannot be effectively separated from single-site damage cases merely using the two PCs, as shown in Fig. 5.8 (b) and (c). Fig. 5.8 (b) shows that the instances belonging to a double-site damage case approximately lie between the clusters formed by the instances belonging to the shared single-site damage cases. This phenomenon indicates that the ICA-based features can distinguish the double-site damage cases from single-site damage cases and represent the correlation between the double- and single-site damage cases if they have shared

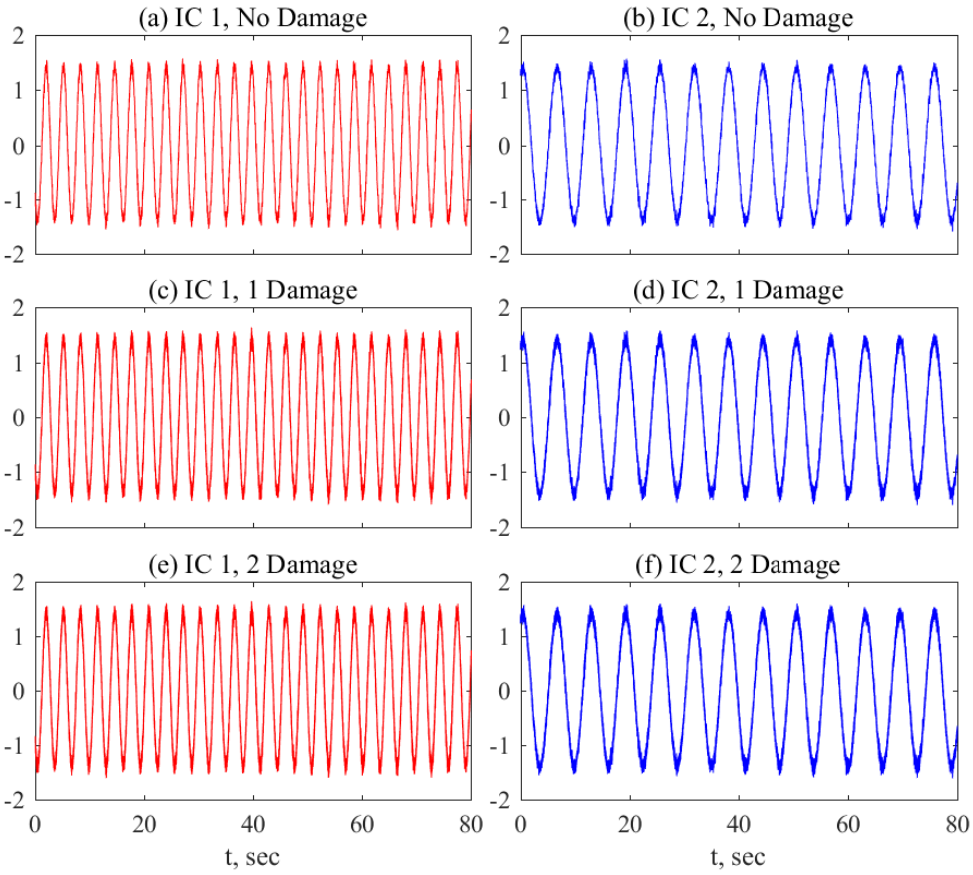


Figure 5.7. Independent components (ICs) from responses of the three damage conditions. (a) and (b) are the ICs of structures without damage; (c) and (d) are the ICs of structures with one damage location; (e) and (f) are the ICs of structures with two damage locations.

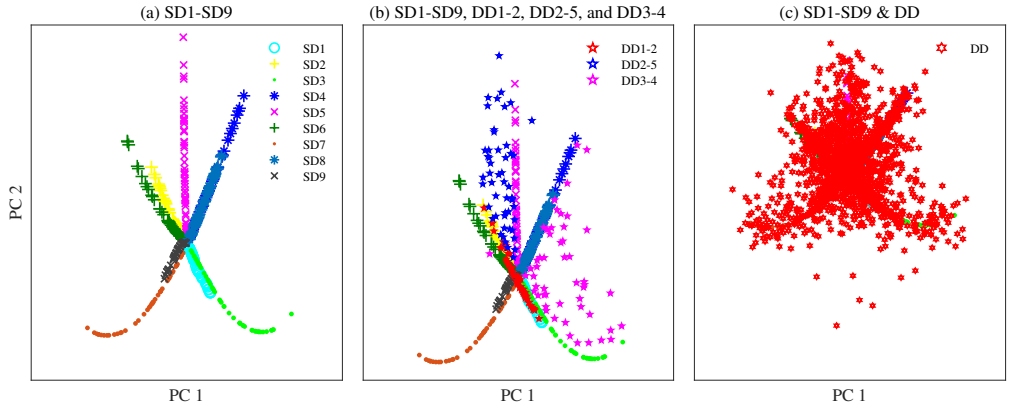


Figure 5.8. Scatter plot of PCA-reduced features. (a) Principal components (PC 1 and PC 2) of single-site damage cases ($\mathrm{SD}_i, i = 1, 2, ..., 9$); (b) PCs of single-site damage cases (SD1 to SD9) and three double-site damage cases (DD1-2, DD2-5, and DD3-4); (c) PCs of single- ($\mathrm{SD}_i, i = 1, 2, ..., 9$) and double- ($\mathrm{DD}i - j, i, j = 1, 2, ..., 9$) damage cases.

damage locations.

To further evaluate the features' performance in identifying the structural damage states, the present study uses the Mahalanobis distance defined as:

$$d_i^{\mathrm{M}} = \sqrt{(\boldsymbol{x}_i - \boldsymbol{\mu})\sum{}^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu})^T} \tag{5.8}$$

where x_i is a sample, $\boldsymbol{\mu}$ is the mean of all the samples belonging to a certain class or cluster, and $\sum$ denotes their covariance matrix. As a measure of dissimilarity between instances, it makes use of the statistical regularities and corrects for the correlation within the data [141, 142]. Therefore, it is advantageous over the commonly used Euclidean distance defined as $d_i^{\mathrm{E}} = \sqrt{(\boldsymbol{x}_i - \boldsymbol{\mu})(\boldsymbol{x}_i - \boldsymbol{\mu})^T}$. The Mahalanobis distance has been widely used for outlier analysis, pattern recognition, and machine learning. Classifiers developed using the Mahalanobis distance have proved to be more accurate than that using the Euclidean distance [142, 143, 144].

Fig. 5.9 shows the distribution of the Mahalanobis distances among the first three single-site damage cases, SD1, SD2, and SD3. One can find that the three diagonal his-

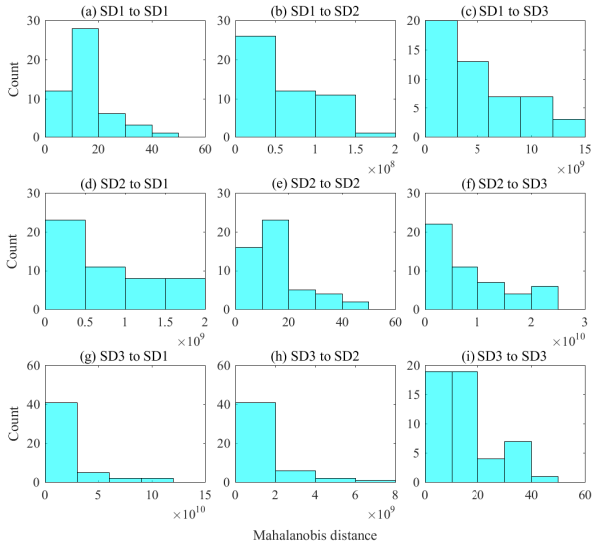


Figure 5.9. Distributions of the Mahalanobis distance between single-site damage cases (SDi and SDj, i, j = 1,2, and 3).

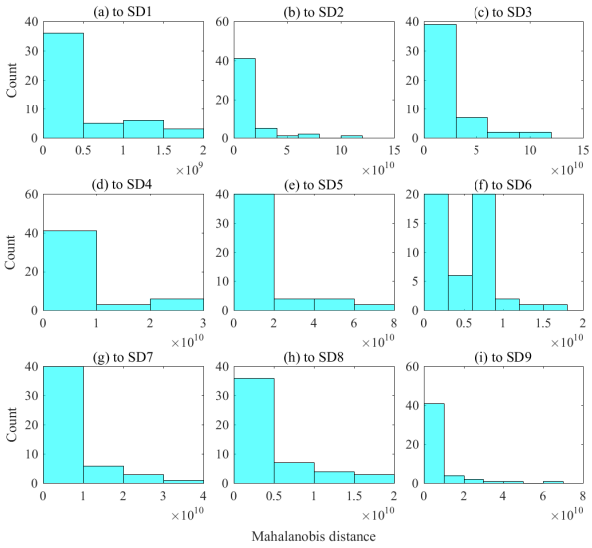


Figure 5.10. Distributions of the Mahalanobis distance between the double-site damage case DD2-7 and the single-site damage cases SD1 to SD9. (a) DD2-7 to SD1; (b) DD2-7 to SD2; etc.

tograms ((a), (e), and (i)) showing the distances of instances to themselves have abscissas bounded within 0 to 60. To the contrary, the off-diagonal histograms representing the distances between different classes (e.g., SD1 to SD2) exhibit distributions in a range of dramatically larger magnitudes from 10^8 to 10^{10}. This phenomenon indicates that the instances of a certain single-site damage state cluster and lie distant from those belonging to other states, when measured with the Mahalanobis distance. Moreover, the performance of the features in representing the similarity within a class and dissimilarity between different classes coincide with the principle of feature extraction and selection, which states that the selected features should minimize the variance within classes and maximize the distances between classes that are measured in the constructed feature space [145]. It is noted that the Mahalanobis distances between other single-site damage cases that are not shown in Fig. 5.9 exhibit similar distributions.

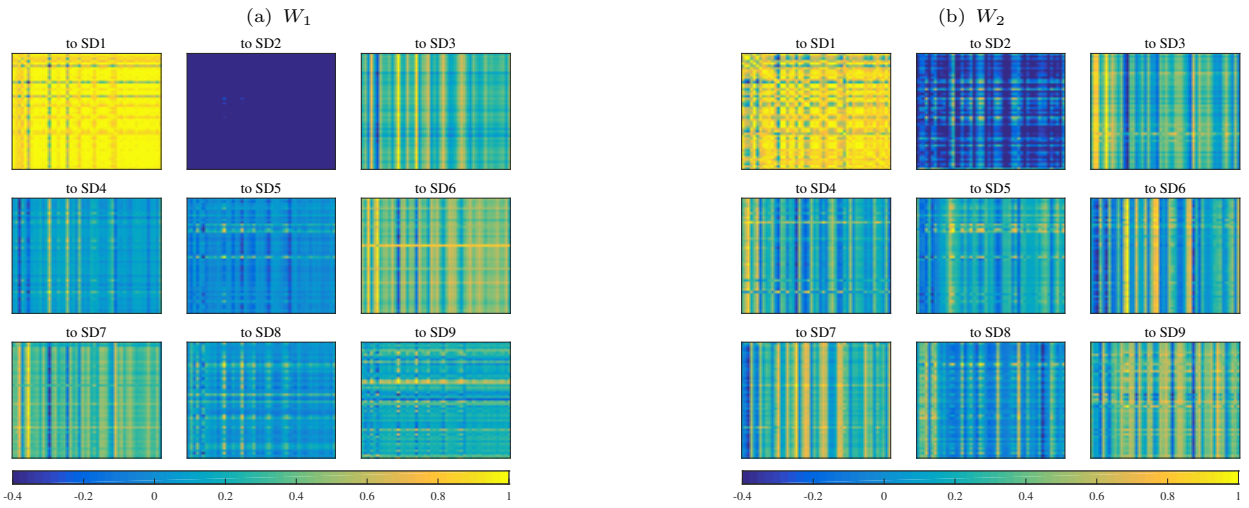


Figure 5.11. Correlation coefficients of single-site damage cases (SD1 to SD$i, i = 1, 2, ..., 9$). The columns correspond to the fifty cases of SD1, and the rows represent the fifty cases of SDi. (a) correlations of W_1, the first column of mixing matrix W; (b) correlations of the second column of mixing matrix, W_2.

Fig. 5.10 shows the distribution of Mahalanobis distances of the double-site damage cases DD2-7(damage at locations 2 and 7) to the nine single-site damage cases, SD1-SD9. One can find that the double-site damage cases, though sharing damage locations with the single-site damage cases SD2 and SD7, do not approach the single site damage cases in the feature space with all the Mahalanobis distances lying on the scale of 10^9 to 10^{10}. The same phenomena can be observed in the other double-site damage cases. Combining the results shown in Fig. 5.9, it can be concluded that the features extracted from the mixing matrix of ICA has the potential to not only identify the single-site damage location but distinguish the cases with double-site damages. Therefore, the proposed methodology can identify the structural damage location(s) progressively using the ICA-based features.

Table 5.3. Error of single-site damage identification using the largest correlation coefficient

Case	1	2	3	4	5	6	7	8	9
error	0.07	0.02	0.00	0.01	0.04	0.00	0.00	0.05	0.22

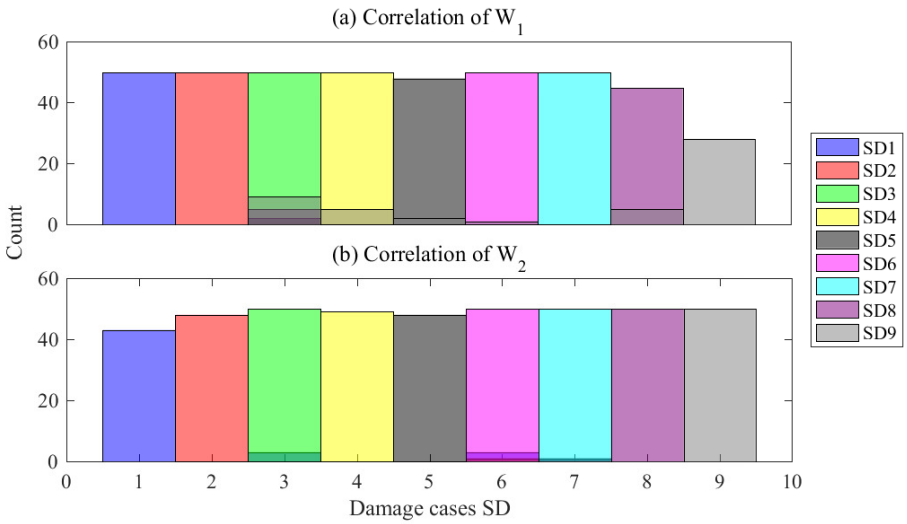


Figure 5.12. Distribution of locations with the largest correlation coefficient (single-site damage cases SD1-SD9).

The present study constructs a classifier, named as M-classifier, based on the Mahalanobis distance and determines the thresholds for decision-making using the Monte Carlo method as described in [146]. The 1% exclusive thresholds are around 150 in this numerical study. If a testing instance has a Mahalanobis distance below the threshold to a certain class representing a single-site damage case, it is determined as belonging to that damage state; otherwise, it is excluded from single-site damage scenario and categorized as a multi-site damage case. Numerical results show that this classifier yields a 100% accuracy when tested on the numerical model in this case study.

5.3.3. Multi-Site Damage Identification

While the multi-site damage cases can be distinguished from the single-site damage cases using the proposed M-classifier, their specific damage locations remain unknown, which will be addressed in this subsection. The mixing matrix element W_{e} represents the contribution of the j^{th} independent component, $S_j(t)$, to the formation of structural

responses at DOF i, $R_i(t)$. The column vector of the mixing matrix W_j incorporates the spatial feature of the independent components $S_j(t)$ and represents how $S_j(t)$ is distributed among the mixtures in ICA [124]. The present study constrains identical output independent components. Then the resulting mixing matrix columns contain the information of how the differing damage states combine and assign the ICs to the measured DOFs. Therefore, the author speculates that the mixing matrix columns from structures sharing damage locations should have larger correlation than otherwise.

Figs. 5.11 (a) and (b) show the correlation coefficients of W_1 and W_2 for the single-site damage cases SD1 to the nine reference single-site damage cases SD1-SD9, respectively. In Fig. 5.11, each subfigure contains 50×50 pixels with each representing a comparison pair. The columns correspond to the cases being analyzed, and the rows represent the reference cases. The bright color (correlation coefficient close to unity) in each figure confirms that the largest correlation mostly happens when comparing two cases with identical

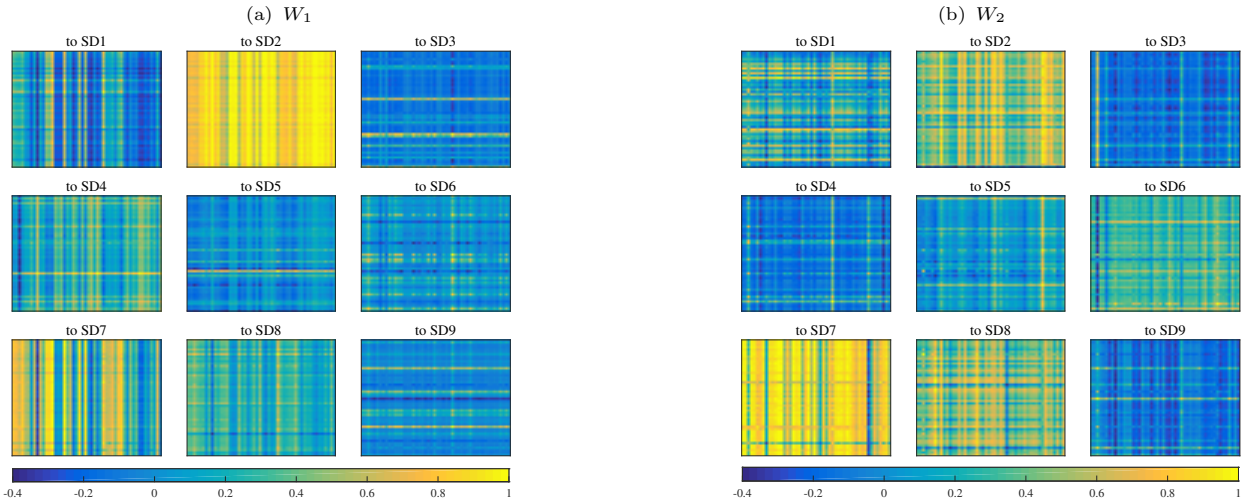


Figure 5.13. Correlation coefficients of double-site damage cases DD2-7 to single -site damage cases SD$i, i = 1, 2, ..., 9$. The columns correspond to the fifty cases of DD2-7, and the rows represent the fifty cases of SDi. (a) correlations of W_1; (b) correlations of W_2.

damage locations. It is noted that similar phenomena occur to other single-site damage cases. Fig. 5.12 shows the distribution of the largest correlation coefficients for the nine single-site damage cases, SD1-SD9. It is found that most of the largest correlation happens when the compared cases have identical damage site. The shaded areas denote the largest correlations occurring elsewhere other than at the target case, which mostly happens to W_1 of SD9 cases. Table 5.3 lists the error rate of identifying single-site damage locations using the largest correlation coefficients. All the cases but SD9 have identification error below 0.10.

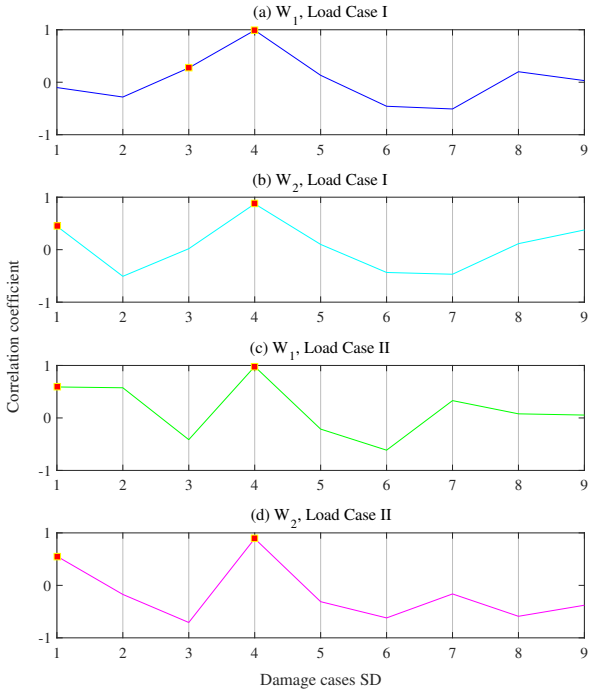


Figure 5.14. Mean correlation coefficients of an example double-site damage case DD1-4 to single-site damage cases SD$i, i = 1, 2, ..., 9$. The square markers on each curve denote the largest two correlations. (a) correlations of W_1 in load case I; (b) correlations of W_2 in load case I; (c) correlations of W_1 in load case II; (d) correlations of W_2 in load case II.

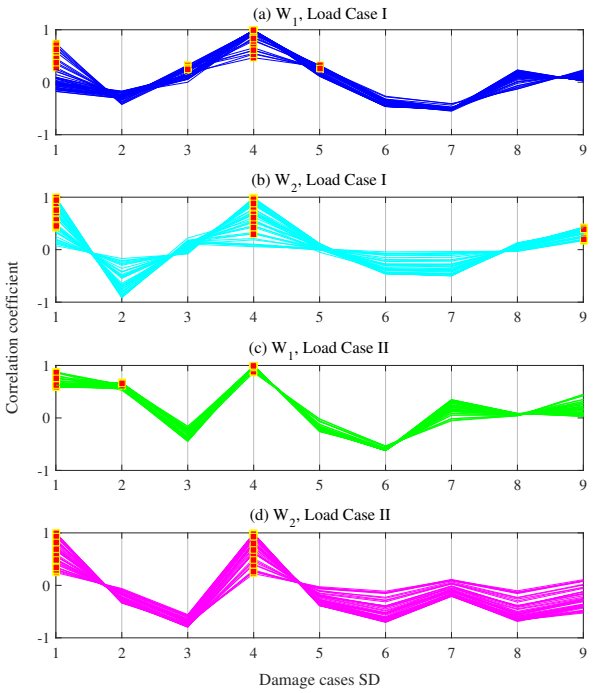


Figure 5.15. Mean correlation coefficients of overall double-site damage case DD1-4 to single-site damage cases SD$i, i = 1, 2, ..., 9$. The square markers on each curve denote the largest two correlations. (a) correlations of W_1 in load case I; (b) correlations of W_2 in load case I; (c) correlations of W_1 in load case II; (d) correlations of W_2 in load case II.

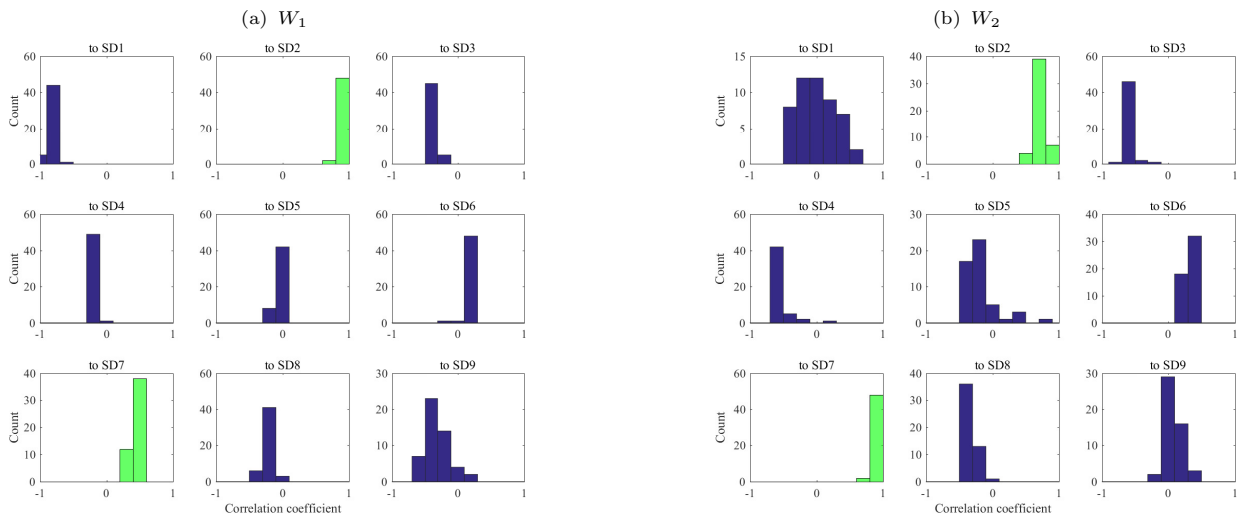


Figure 5.16. Distribution of correlation coefficients of an example double-site damage case DD2-7 to single-site damage cases SD$i, i = 1, 2, ...9$. (a) correlations of W_1; (b) correlations of W_2.

Similar to Fig. 5.11 for single-site damage cases, Fig. 5.13 plots the correlation coefficients of a double-site damage case, DD2-7 with damages at Locations 2 and 7, to the nine single-site damage cases. As can be observed, W_1 and W_2 of DD2-7 cases exhibit large correlation with that of SD2 and SD7, each sharing a damage location with DD2-7. Via comparing Figs. 5.13 (a) and (b), one can find that the results from two columns of mixing matrix are mutually supplementary for double-site damage identification: W_1 yields the largest correlation coefficients when comparing DD2-7 with SD2, and W_2 produces the largest correlation when DD2-7 is compared with SD7. Figs. 5.16 presents the distribution of correlation coefficients of one case of DD2-7 to single-site damage cases SD1-SD9. It shows that the DD2-7 case exhibits apparently larger correlation to SD2 and SD7 cases than to others. In detail, W_1 has the largest correlation coefficients to SD2 that are all above 0.6; W_2 yields the largest correlation with SD7 that are also above 0.6.

It should be noted that some double-site damage cases do not show such obvious

correlation with the two target single-site damage cases as exhibited by the DD2-7 case. Numerical analysis manifests that this drawback can be addressed via using additional load case in the dynamic tests. Fig. 5.14 shows the mean correlation coefficients of W_1 and W_2 of one example case of DD1-4 to the single-site damage cases SD1-SD9 using the responses under two load cases (see Table 5.2 for the load information). The red squares on the curves denote the two SD cases that yield the largest correlation to this example DD1-4 case. Figs. 5.14 (a) and (b) show that for load case I, the largest correlation of W_1 occurs at SD3 and SD4 and W_2 has the largest correlation at SD1 and SD4, which recommends only Location 4 as the damage site. However, Figs. 5.14 (c) and (d) show that both the correlations of W_1 and W_2 from load case II suggest both Location 1 and Location 4, which are identical to the target damage condition. Therefore, if combining the results using load cases I and II and picking the two most commonly recommended locations using the largest correlation coefficients, the damage locations of this DD1-4 case can be correctly identified. The results of all the fifty DD1-4 cases as shown in Fig. 5.15 verifies the efficiency of this method.

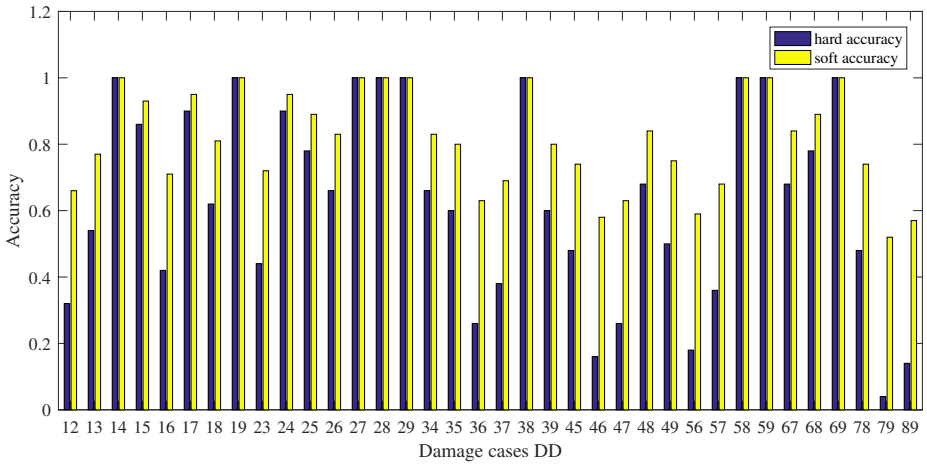


Figure 5.17. Identification accuracy of double-site damage cases (DD1-2 to DD8-9)

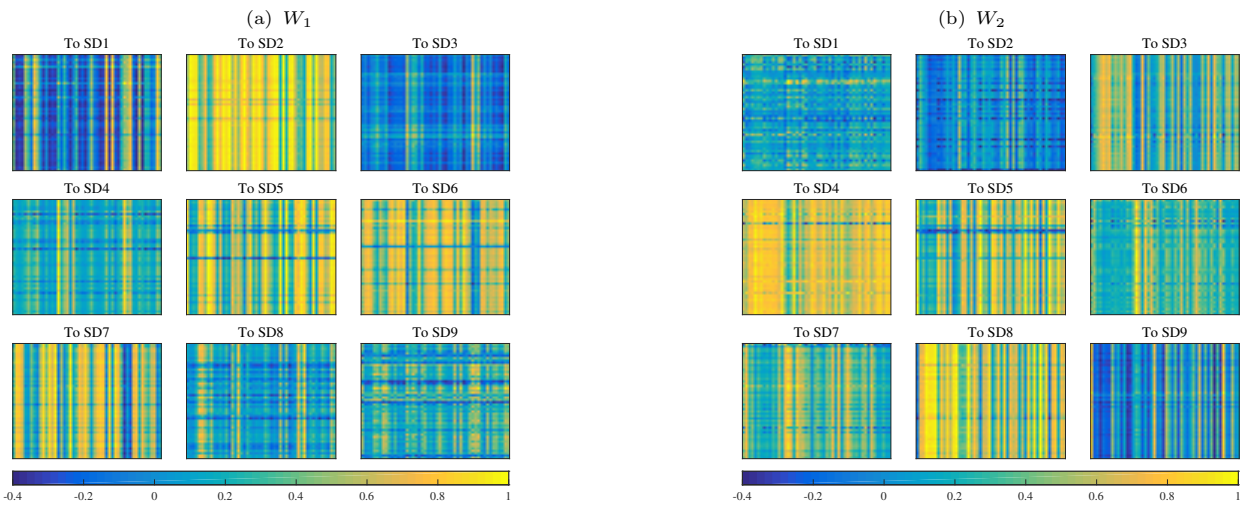


Figure 5.18. Correlation coefficients of three-site damage cases TD2-5-8 to single -site damage cases SD$i, i = 1, 2, ..., 9$. The columns correspond to the fifty cases of TD2-5-8, and the rows represent the fifty cases of SDi. (a) correlations of W_1; (b) correlations of W_2.

For many of the double-site damage cases that cannot be correctly identified, the method proposed above still have its value. For example, one correct location is identified though not both of them, or the target locations are ranked the 2^{nd} and 3^{nd} instead of the first two. These imperfect results still provide valuable suggestions for structural damage identification. Fig. 5.17 lists the identification accuracies of the double-site damage cases DD12-DD89. Hard accuracy counts only the identically localized damage cases, and soft accuracy counts half for the imperfectly identified cases mentioned above. The averages are 0.69 for the hard accuracy and 0.82 for the soft accuracy. The evident difference between the hard accuracy and soft accuracy for many cases confirms that the proposed method generates informative, though not necessarily complete, suggestions for double-site structural damage identification.

In addition to double-site damage identification, performance of the proposed method is also examined on a three-site damage case, TD2-5-8 (damage at locations 2,

5, and 8). Fig. 5.18 illustrates the correlation coefficients of W_1 and W_2 to SD1 to SD9. W_1 exhibits large correlation to SD2, SD5, SD6, and SD7, and W_2 yields large correlation to SD3, SD4, SD5, and SD8, as indicated by the color difference. On the other hand, inspecting the plots column-wise, one can find that the largest correlation of W_1 happens most to SD2 and SD5 and the largest correlation of W_2 happens most to SD5 and SD8. Fig. 5.19 shows the mean correlation coefficients of one example case TD2-5-8. It shows that both W_1 and W_2 recommend correct damage locations in both dynamic tests. Fig. 5.20 further verifies this encouraging finding with the correlation coefficients of all the fifty TD2-5-8 cases.

5.4. Summary

This study attempts to identify structural damage at single and multiple sites using structural responses. The major difficulty of locating multi-site damage lies in the confusion from damage severity which invalidates most vibration-based damage features that work well for single-site damage detection. The present paper proposes a methodology for damage detection using the constrained independent component analysis (cICA). The cICA in this paper is realized through providing an initial mixing matrix to ICA and updating it iteratively until the output independent components converge. The benefit of cICA is that it enforces the constraint of identical independent components (ICs) and thus compacts all the valuable information related to structural damage into the mixing matrix. Taking the mixing matrix elements as features, an M-classifier based on the Mahalanobis distance is introduced. The M-classifier can accurately identify the single-site

damage and distinguish the multi-damage cases. Subsequently, the multi-damage identification is achieved by evaluating the correlation of the mixing matrix columns with that of the reference single-site damage cases. The outcome shows that the proposed method can provide reliable judgment for determining the multiple damage sites. The major contribution of this study includes: it proposes a novel method of extracting compact damage-sensitive features using cICA, which works efficiently for both single- and multi-site structural damage identification; the extracted features can be used as input to most classifiers in machine learning for damage detection; it proposes a damage identification method using the cICA outputs that progressively identifies single- and multi-site structural damage, which has the potential to largely improve the damage detection efficiency; the proposed method has the potential to relieve the burden of data-driven SHM from data incompleteness/insufficiency for pattern recognition. Future work will validate the proposed methodology using experimental studies and field testing on real structures via applying multiple types of loading, including harmonic excitations, impact loading, and environmental excitations, so as to prove the practical feasibility and generality of the proposed methodology.

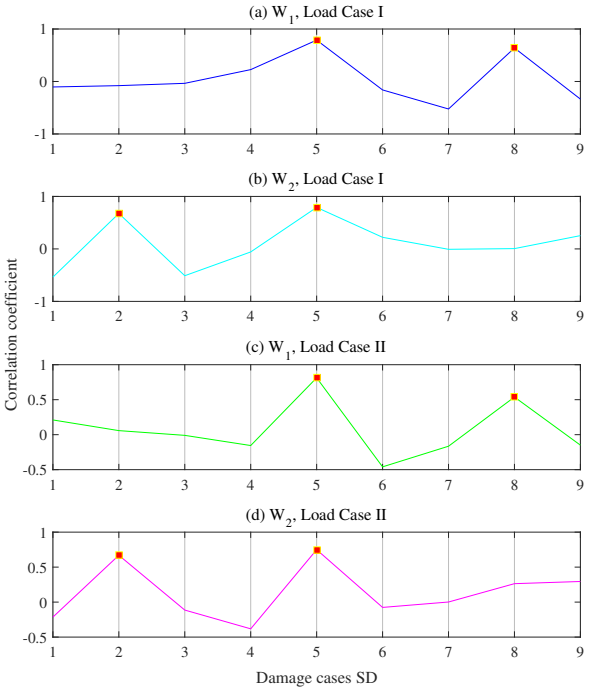


Figure 5.19. Mean correlation coefficients of an example three-site damage case TD2-5-8 to single-site damage cases SD$i, i = 1, 2, ..., 9$. The square markers on each curve denote the largest two correlations. (a) correlations of W_1 in load case I; (b) correlations of W_2 in load case I; (c) correlations of W_1 in load case II;(d) correlations of W_2 in load case II.

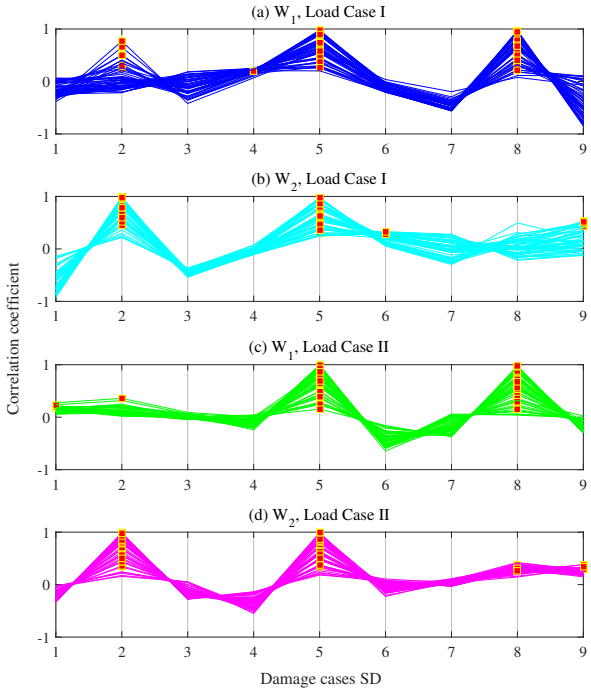


Figure 5.20. Mean correlation coefficients of overall three-site damage case TD2-5-8 to single-site damage cases SD$i, i = 1, 2, ..., 9$. The square markers on each curve denote the largest two correlations. (a) correlations of W_1 in load case I; (b) correlations of W_2 in load case I; (c) correlations of W_1 in load case II;(d) correlations of W_2 in load case II.

Chapter 6. Structural Damage Identification of Offshore Wind Turbines: a Two-Step Strategy via FE Model Updating

6.1. Introduction

In Chapters 2 and 3, physics-guided machine learning and transfer learning methods are investigated respectively to integrate data-driven and model based SHM methods for improved performance of damage identification; Chapters 4 and 5 present two methods for data-driven multi-site damage detection on condition that insufficient data are available for learning a robust statistical pattern. The present chapter presents a case study of structural damage identification on offshore wind turbines. A two-step strategy of damage evaluation is developed based on FE model updating.

Offshore wind energy production has increased significantly in recent years. According to the Global Wind Report 2018 [147], more portions of offshore wind turbines (OWTs) are installed in the past years: from 5.3% in 2015 to 8.8% in 2018. The past and projected future increase of OWTs poses a severe challenge to the operation and maintenance (O&M). O&M is costly for WTs considering the height of a wind turbine tower, the blade size, and the locations of wind turbines at remote mountainous or marine regions. For example, a 750 kW turbine with a 20-year expected service life has an O&M cost of 25% to 30% of the overall energy generation cost or 75–90% of the investment costs [148]. Regarding OWTs installed far from the coast, the challenge is further increased due to the extreme weather conditions and complex dynamics, which affects technical details such as sensor tolerance and endurance, data acquisition and transmission, etc. [149, 150, 151]. As a result, The O&M costs account for a substantial part of the OWTs' total life costs es-

pecially if the wind farms are deployed in deeper waters and harsher marine environments [152].

With a potential of reducing costs and improving convenience of OWTs' O&M, structural health monitoring (SHM) has attracted an interest from the OWT community and is expected to be more essential for OWTs in the next several decades [153]. SHM is an emerging research topic for OWTs due to the challenges that offshore environment poses and the associated costs for installation and data processing. Currently, SHM of OWTs is slowly entering the field deployment stage and becoming much in demand for OWTs with the increasing power rate and distance of installation from the coast [153]. SHM has the potential of avoiding unscheduled maintenance, identifying design weaknesses before failure, eliminating unexpected catastrophic failures, improving the availability of power production while preventing wind turbines' overloading, allowing less conservative blade design for capturing more wind energy, and assessing remaining life of OWT structural components [153, 156, 157, 158]. Economic benefits of SHM implementation in the support structure of OWT have been proved in [159] by comparing the increased capital expenditure (CAPEX) with the decreased operational expenditure (OPEX). In summary, SHM is promising to improve OWT's profitability, reliability, and sustainability through systematically designed hardware installation, data acquisition and processing, and decision making.

However, SHM for OWTs faces multiple challenges in practical application. First, compared with the SHM technologies applied to civil or oil and gas industries, the SHM system for OWTs is subject to limited accessibility, severe environmental effects, and operational loads. Hence, the SHM system for OWTs needs to be specially designed regard-

ing hardware selection and installation, sensor deployment and data acquisition, etc. for a certain OWT considering limited examples to follow according to a survey in [153]. Currently, there is no widely accepted practice regrading the specification of OWT monitoring systems, as industry makes every wind farm unique with respect to implemented technologies, sensor location and quantities, redundancies, etc. [160]. Second, operational modal analysis (OMA) that is widely used in existing studies of WTs can be substantially challenging for OWTs as reported in [161, 162, 163, 164, 165]. The major challenges include the time variance of the OWT system, the violation of the steady state random excitation assumptions, selection of reference frame, and mathematical uncertainties. These challenges cause remarkable scatter in the extracted modal properties, i.e., natural frequencies, mode shapes, and modal dampings, which limits the application of the OMA outcome in assessing the structural conditions of OWTs.

Despite the challenges mentioned above, a variety of SHM techniques are used on OWTs including acoustic emission [166] and ultrasonic testing [167] for monitoring cracking, debonding, etc. on blades, modal properties monitoring to detect scour of foundation [168], and strain monitoring for continuous operational state monitoring at critical locations of OWT [169]. However, many advanced SHM methods and algorithms such as FE model updating that are extensively studied and widely applied in civil and mechanical engineering are rarely used in SHM of OWTs. In addition to the challenge mentioned above regarding system installation, data acquisition, signal processing for example in OMA, another difficulty lies in the complex modeling of the wind turbine system especially in operation.

Though rarely used in the SHM of OWT, FE model updating using vibration based

features (e.g., modal properties) has been widely applied for structural identification and damage detection of engineering structures [37, 38, 39]. To this end, modal properties including natural frequencies and mode shapes are usually obtained through operational modal analysis from ambient vibrations or from forced vibration tests. FE model updating has the potential of identifying the structural damage occurrence, damage locations, and its severities from the variation of identified structural mechanical parameters [40]. Examples of FE model updating used for OWT monitoring can be found in [170, 171, 172, 173, 174]. The target of model updating in these studies is either a numerical model of a certain OWT compoent (blade in [170, 173], tower in [171], support [174]) or a simlified OWT model [172] without considering the operational conditions.

The present study attempts to improve the performance of structural damage assessment on OWTs through applying FE model updating to two numerical OWT models at different scales. The proposed methodology first identifies the probable damaged structural component(s) on an OWT, i.e. blade(s) and/or tower using a verified global mathematical OWT model which was developed for vibration control [175, 176, 177, 178, 179]. With the identified damaged component(s), the second step further localizes and quantifies the structural damage through updating a detailed FE model. It circumvents the requirement of intensive computation due to the complexity of OWT models in SHM and establishes a framework of a two-step strategy of damage identification via FE model updating.

6.2. Methodology: a Two-Step Strategy via FE Model Updating

6.2.1. FE Model Updating

For a monitored structure, the stiffness matrix $\mathbf{K}$ can be formulated as:

$$\mathbf{K} = \mathbf{K}_0 + \sum_{i=1}^{n_\alpha} \alpha_i \mathbf{K}_i \tag{6.1}$$

in which $\mathbf{K}_0$ is the sum of reference substructural stiffness matrices prior to model updating, which usually corresponds to an intact structural condition; $\mathbf{K}_i$ is the nominal stiffness matrix of substructure i with unknown stiffness; α_i is the coefficient corresponding to $\mathbf{K}_i$; n_α is the number of substructures with unknown stiffness. Hence, $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_{n_\alpha}]$ containing all unknown stiffness coefficients is the target of FE model updating. The identified variation in $\boldsymbol{\alpha}$ during model updating reflects the degradation of certain structural components or substructures.

In this study, deterministic model updating (DMU) is implemented through minimizing an objective function that represents the discrepancy between the numerical model and the tested structure. In this study, the objective function in DMU is formulated using the difference in eigen-frequency, mode shape, and mode shape curvature between the measurement and estimation from numerical simulation. That is:

$$\begin{aligned} &L(\boldsymbol{\alpha}) = \\ &W_\lambda \left[\sum_{i=1}^{n_\mathrm{m}} \left(\frac{\lambda_i^\mathrm{e} - \lambda_i^\mathrm{m}(\boldsymbol{\alpha})}{\lambda_i^\mathrm{e}} \cdot w_{\lambda_i} \right)^2 \right] + \\ &W_\Phi \left[\sum_{i=1}^{n_\mathrm{m}} ||(\boldsymbol{\Phi}_i^\mathrm{e} - \boldsymbol{\Phi}_i^\mathrm{m}(\boldsymbol{\alpha})) \cdot w_{\Phi_i}||_2^2 \right] + \\ &W_{\Phi''} \left[\sum_{i=1}^{n_\mathrm{m}} \left|\left| \left(\boldsymbol{\Phi}_i^{''\mathrm{e}} - \boldsymbol{\Phi}_i^{''\mathrm{m}}(\boldsymbol{\alpha}) \right) \cdot w_{\Phi_i''} \right|\right|_2^2 \right] \end{aligned} \tag{6.2}$$

in which n_{m} is the number of measured modes in dynamic tests; λ_i^{e} is the experimentally measured eigen-frequency of the i^{th} mode; $\lambda_i^{\mathrm{m}}(\boldsymbol{\alpha})$ is the evaluated value of λ_i from the FE model using a certain value of $\boldsymbol{\alpha}$; $\boldsymbol{\Phi}_i^{\mathrm{e}}$ is the measured mode shape of the i^{th} mode; $\boldsymbol{\Phi}_i^{\mathrm{m}}(\boldsymbol{\alpha})$ is the evaluated $\boldsymbol{\Phi}_i$ at the measured DOFs using $\boldsymbol{\alpha}$; $\boldsymbol{\Phi}_i^{''\mathrm{e}}$ is the measured mode shape curvature of the i^{th} mode; $\boldsymbol{\Phi}_i^{''\mathrm{m}}(\boldsymbol{\alpha})$ is the evaluated $\boldsymbol{\Phi}_i^{''}$ at the measured DOFs using $\boldsymbol{\alpha}$; w_{λ_i}, w_{Φ_i}, and $w_{\Phi_i''}$ are the inner weighting factors of eigen-frequency, mode shape, and mode shape curvature of the i^{th} mode respectively. The inner weight factors balance the contribution from each mode based on their magnitude, measurement reliability, etc. In comparison, W_{λ_i}, W_{Φ_i}, and $W_{\Phi_i''}$ are the outer weighting factors balancing the overall sum of eigen-frequency, mode shape, and mode shape curvature terms.

With the optimization problem formulated for DMU, the Levenberg-Marquardt algorithm is selected for solving and is implemented using the "lsqnonlin" solver in MATLAB. Following the practice in [180], a number of runs, for example 100, are implemented with random starting points, and the solution yielding the least objective function is selected as the final solution. With the updated structural parameters $\boldsymbol{\alpha}$ of a certain state, the possible damage location(s) and quantities can be evaluated by comparing the value of $\boldsymbol{\alpha}$ with that estimated from a reference or intact structure, that is $\boldsymbol{\alpha}_0$.

6.2.2. Damage Introduction to the Offshore Wind Turbine

This section defines how structural damage is introduced to the OWT in numerical simulations of this study. Figures 6.1 (a) and (b) show OWT blades and tower, respectively. Segment number is denoted increasingly from the root to the tip of the blade and from the bottom to the top of the tower. Damage is introduced via reducing the stiffness of certain segment(s) of one or more component(s) uniformly by 10% to 50%. When

damage happens to certain segment(s) of certain component(s), the stiffness terms corresponding to the damaged components in the stiffness matrix of the analytical model (see Section 3.1) will decrease accordingly. In step 2, damage is directly introduced by reducing the modulus of elasticity of elements contained in the damaged segments in the established ANSYS FE OWT model (see Section 4.1).

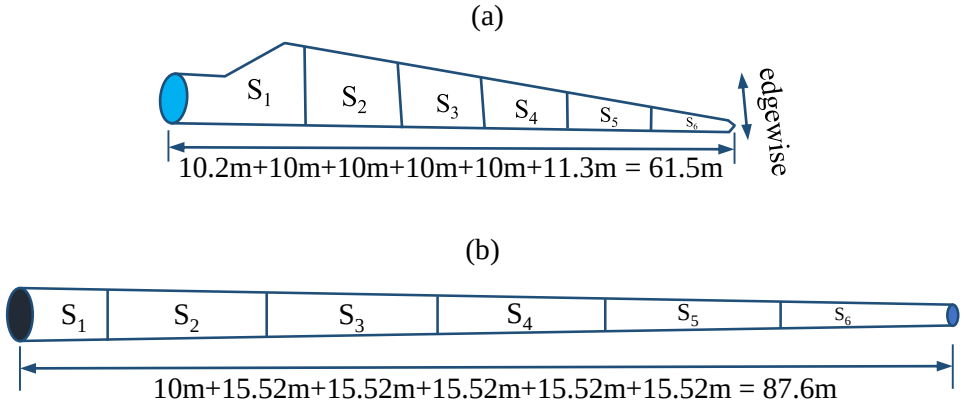


Figure 6.1. Segmentation of (a) blade and (b) tower in damage detection.

6.2.3. A Two-step Damage Identification Strategy

This study proposes a two-step strategy of identifying structural damage on OWT through model updating. This section generally describes the framework of this strategy, and more details can be found in Sections 3 and 4. Considering the complexity of an OWT model, it is challenging to identify structural damage through updating a full-scale detailed FE model. To overcome this challenge, damage identification in the present study is divided into two steps: (1) identifying the damaged component(s) (blades, tower, or foundation) using a verified global OWT model under operational conditions; (2) more accurately localizing the quantifying damage on the identified damaged component(s) via updating a detailed substructural model. The damaged component in step 1 can be a certain blade, the tower, or the foundation. This study mainly focuses on identifying damage

on blades and tower. The global mathematical model in step 1 is an analytical model established considering the fundamental mode of each component using the Euler-Lagrange equation. The target of parameter updating of DMU in step 1 is set as the selected diagonal elements in the derived stiffness matrix of the analytical model which are directly related to the global damage of OWT components. Hence, the variation of the updated structural parameter values can reflect damage occurrence on certain component(s). In step 2, a detailed FE model of the 5 MW OWT is established for numerical simulation using ANSYS. Moreover, to achieve efficient DMU, substructural FE models are built for each of the damaged component(s) identified in step 1. This approach of model reduction largely decreases the dimension of the optimization problem in DMU and thus significantly increases the efficiency and accuracy of damage identification. Model updating with respect to the substructural model yields damage locations and severities of each damaged component.

Figure 6.2 illustrates the general framework of damage identification of OWTs through DMU in a two-step strategy proposed in this study. Given the OWT model, namely the generalized model in step 1 or the detailed FE model in step 2, numerical simulation yields the dynamic responses of the OWT model under simulated wind and wave loads. With the simulated responses, OMA via frequency domain decomposition (FDD) [181] generates the modal properties that are necessitated for formulating the objective function in Eqn. (6.2). In this study, only accelerations are used in OMA considering the difficulty of measuring displacement and velocity responses in reality. Additionally, mode shape curvature is not involved in step 1 because of the limited DOFs considered

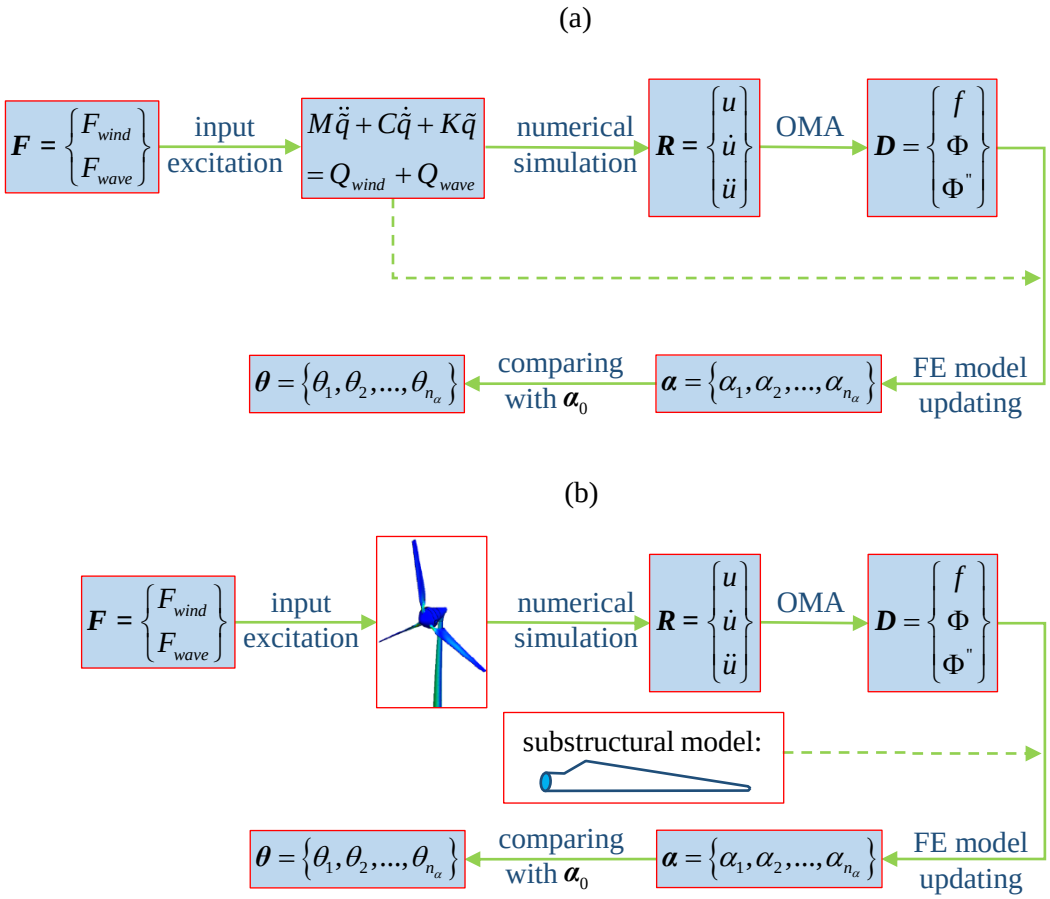


Figure 6.2. Framework of OWT damage identification via DMU using two-step strategy. (a) framework of step 1; (b) framework of step 2. $\boldsymbol{F}$ is the load vector containing the wind and wave loads, that is F_{wind} and F_{wave}. The equation of motions in (a) is based on the analytical model in step 1 and will be derived in Section 6.3. $\boldsymbol{R}$ denotes the dynamic responses containing displacement u, velocity $\dot{u}$, and acceleration $\ddot{u}$. OMA stands for operational modal analysis. $\boldsymbol{D}$ denotes the output of OMA including the natural frequency f, mode shape Φ, and mode shape curvature Φ''. $\boldsymbol{\alpha}$ contains updated structural parameter values for a certain state; $\boldsymbol{\alpha}_0$ is the parameter vector for a reference or intact structure. $\boldsymbol{\theta}$ is the identified damage severities at possible damaged locations. The dashed arrow indicates that a reference OWT model is required in model updating, which can be the global mathematical model in step 1 or the substructural model in step 2.

in the global mathematical model. Hence, the corresponding objective function in Eqn. (6.2) does not contain the term regarding mode shape curvature. Then DMU is conducted through optimizing the structural parameter $\boldsymbol{\alpha}$ towards minimizing the objective function, which yields the updated $\boldsymbol{\alpha}$. In this process, the outer weighting factors are tuned to

balance the contribution from each term. Finally, the damage locations and severities can be obtained by comparing $\boldsymbol{\alpha}$ with that of the reference or intact structure $\boldsymbol{\alpha}_0$.

6.3. Damaged Structural Component Identification

6.3.1. Establishment of the Global Mathematical Model

A schematic model of an OWT under wind and wave loading is demonstrated in Fig. 6.3 where the origin of the global coordinate system is set at the mean sea level. The OWT model presented in this chapter has 12 degree of freedom (DOF). Variables q_1 to q_3 are the coordinates of the three blades in edgewise direction, and q_4 to q_6 are the coordinates of the three blades in flapwise direction. The nacelle fore-aft and side-side relative coordinates with respect to the foundation are represented by q_7 and q_8. Variables q_9 to q_{12} are the foundation coordinates in the translational and rotational directions. Constraints of soil are modeled using spring with coefficients of k_x, k_y, $k_{x\phi}$ and $k_{y\phi}$. The soil damping effect is modeled using dash-pot dampers with coefficients of c_x, c_y, $c_{x\phi}$ and $c_{y\phi}$.

In the present study, the mathematical model of the OWT is established using the Euler-Lagrangian equation:

$$\frac{d}{dt}\frac{\partial T(t,\tilde{q}(t),\dot{\tilde{q}}(t))}{\partial \dot{q}_i(t)} - \frac{\partial T(t,\tilde{q}(t),\dot{\tilde{q}}(t))}{\partial q_i(t)} + \frac{\partial V(t,\tilde{q}(t))}{\partial q_i(t)} = Q_i(t) \tag{6.3}$$

where T and V denote the OWT kinetic and potential energy; $\tilde{q}(t)$ denotes the generalized coordinate vector; $Q_i(t)$ is the generalized force corresponding to $q_i(t)$.

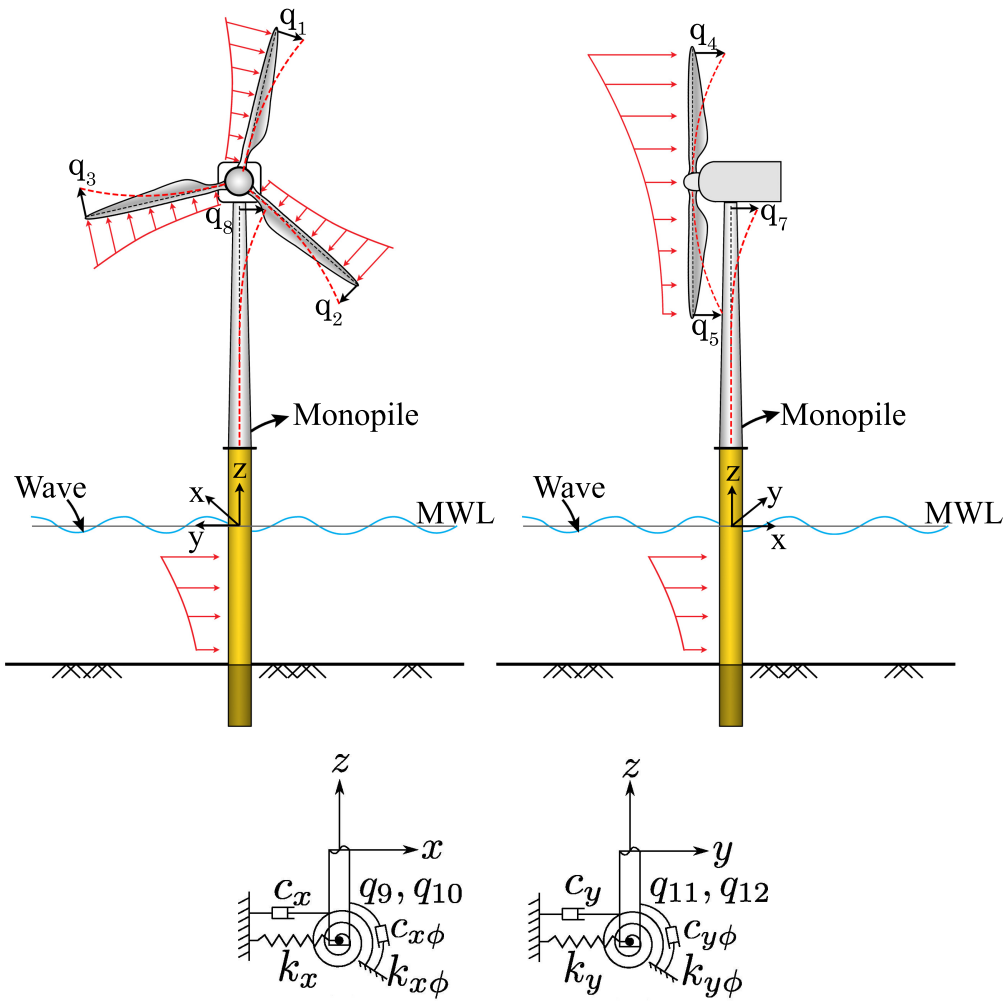


Figure 6.3. Schematic model of the monopile OWT. q_1 to q_3 are the DOFs of blades in edgewise direction; q_4 to q_6 are the DOFs of blades in flapwise direction; q_7 and q_8 are the DOFs of nacelle in the fore-aft and side-side direction respectively; q_9 to q_{12} are the foundation DOFs; k_x, k_y, $k_{x\phi}$ and $k_{y\phi}$ are the foundation spring constants; c_x, c_y, $c_{x\phi}$ and $c_{y\phi}$ are the foundation damping coefficients.

The displacements of the nacelle can be determined as:

$$\begin{aligned} u_{nac}^{fa} &= q_7 + q_9 + h\tan(q_{10}) \approx q_7 + q_9 + hq_{10} \\ u_{nac}^{ss} &= q_8 + q_{11} + h\tan(q_{12}) \approx q_8 + q_{11} + hq_{12} \end{aligned} \tag{6.4}$$

where u_{nac}^{fa} and u_{nac}^{ss} denote the nacelle displacement in the fore-aft and side-side directions, respectively; h is the nacelle height.

The velocity components of the nacelle can be expressed as:

$$\begin{aligned} v_{nac}^{fa} &= \dot{q}_7 + \dot{q}_9 + h\dot{q}_{10} \\ v_{nac}^{ss} &= \dot{q}_8 + \dot{q}_{11} + h\dot{q}_{12} \end{aligned} \tag{6.5}$$

where v_{nac}^{fa} and v_{nac}^{ss} denote the nacelle velocity in fore-aft and side-side directions.

Then the total kinetic energy of the OWT system can be expressed as:

$$\begin{aligned} T = \frac{1}{2}\sum_{j=1}^{3}\int_0^R \overline{m} v_{bj}^2(r,t)dr + \frac{1}{2}M_{nac}v_{nac}^2 + \frac{1}{2}\int_0^h \overline{M} v_{tow}^2 dz \\ + \frac{1}{2}M_f[\dot{q}_9^2(t) + \dot{q}_{11}^2(t)] + \frac{1}{2}I_f[\dot{q}_{10}^2(t) + \dot{q}_{12}^2(t)] \end{aligned} \tag{6.6}$$

where M_{nac} is the mass of the nacelle; M_f and I_f are the mass and moment of inertia of the foundation; v_{bj} and v_{tow} are the velocity of the blade and the tower which is detailed in [176]; $\overline{m}$ and $\overline{M}$ are the blade and tower mass density per unit length.

The potential energy V of the OWT is expressed as:

$$V = V_b + \frac{1}{2}k_t^{fa}q_7^2(t) + \frac{1}{2}k_t^{ss}q_8^2(t) + \frac{1}{2}k_x q_9^2(t) + \frac{1}{2}k_y q_{11}^2(t) + \frac{1}{2}k_{x\phi}q_{10}^2(t) + \frac{1}{2}k_{y\phi}q_{12}^2(t) \tag{6.7}$$

where V_b denotes the potential energy of the blades which is detailed in [176], k_t^{fa} and k_t^{ss} are the tower stiffness in fore-aft and side-side directions.

The equation of motion of the OWT can be derived through substituting Eqn. (6.6) and Eqn. (6.7) into Eqn. (6.3).

$$\widetilde{M}\ddot{\tilde{q}} + \widetilde{C}\dot{\tilde{q}} + \widetilde{K}\tilde{q} = \widetilde{Q}_{wind} + \widetilde{Q}_{wave} \tag{6.8}$$

where parameters $\widetilde{M}$,$\widetilde{C}$ and $\widetilde{K}$ denote the mass, damping and stiffness matrices; $\widetilde{Q}_{wind}$ and $\widetilde{Q}_{wave}$ are the generalized force vectors of wind and wave loading. Details of $\widetilde{M}$,$\widetilde{C}$ and $\widetilde{K}$ and the derivation of $\widetilde{Q}_{wind}$ and $\widetilde{Q}_{wave}$ can be found in [176, 177].

6.3.2. Operational Modal Analysis

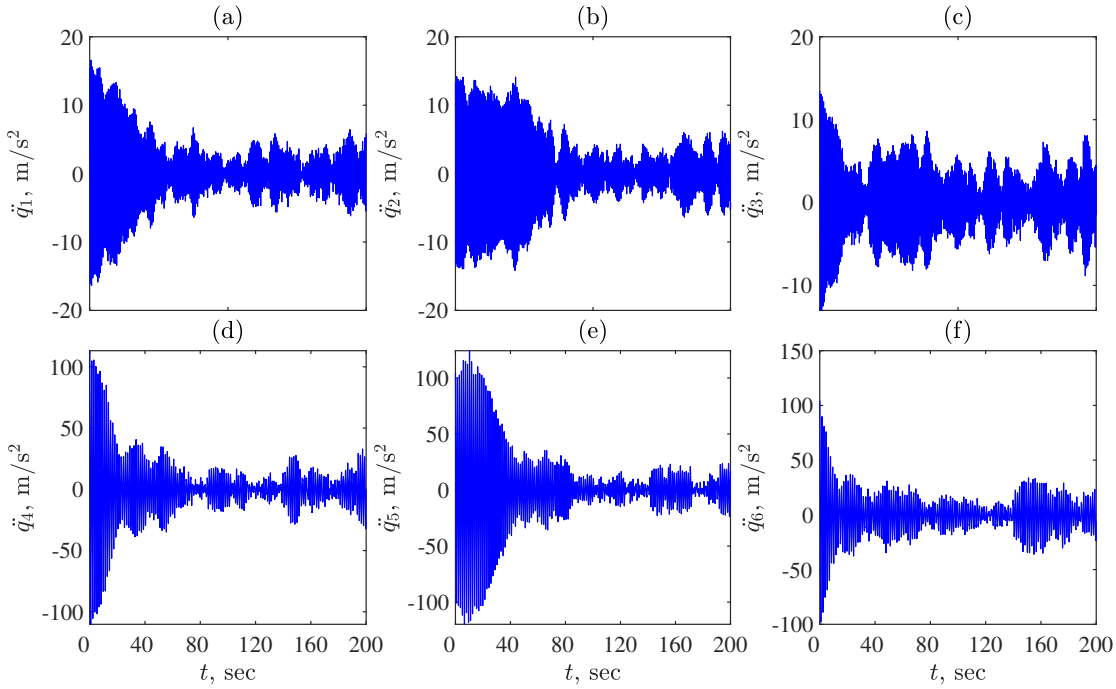


Figure 6.4. Structural responses of wind turbine under intact condition.

Solving the equations of motion in Eqn. (6.8) yields the responses of the OWT in terms of the defined DOFs. Figures 6.4 (a) to (f) show the accelerations of blade tips corresponding to the DOFs defined in Figure 6.3. Accelerations are used to extract modal properties through OMA. Figure 6.5 compares the OMA outputs for the intact condition and a damaged condition. The damaged condition in this example is a general damage case with significant stiffness reduction on each OWT component as shown in the last column of Table 6.1. The DOFs q_9 to q_{12} corresponding to foundation motions are not shown in mode shapes (see Figure 6.5 (b) and (c)) considering their limited magnitudes. Figure 6.5 shows that the OMA outputs of the intact condition are very close to that of the reference values, approving the quality of system identification via OMA using only measured acceleration responses. Additionally, comparing the OMA results of the intact structure

with that of the damaged structure shows that structural damage (stiffness reduction) causes considerable frequency reduction (see Figure 6.5 (a)) and mode shape variations (see Figure 6.5 (b)). It can be expected that DMU with the extracted modal properties from the two conditions will yield structural parameters with recognizable difference, indicating damage occurrence at certain locations.

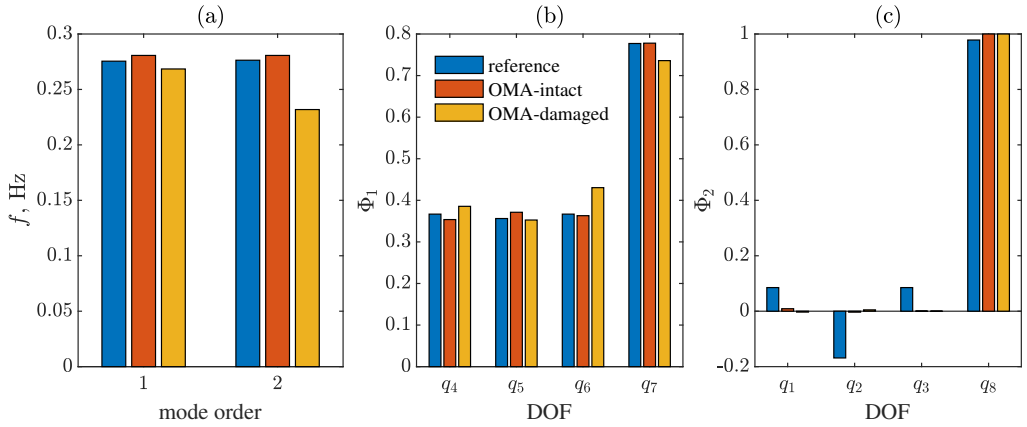


Figure 6.5. Comparison of modal properties from OMA and the reference. The reference properties are calculated using the reference structural matrices $\mathbf{M}_0$ and $\mathbf{K}_0$. In (b), mode 1 (Φ_1) denotes the first out-of-plane mode involving DOFs 4-7, and mode 2 (Φ_1 in (c)) denotes the first in-plane mode involving DOFs 1-3 and 8. (a) frequencies; (b) shape of mode 1; (c) shape of mode 2.

6.3.3. FE Model Updating and Structural Damage Evaluation

Model updating is implemented by solving the optimization problem with the objective function formulated in Eqn. (6.2) using the extracted frequencies and mode shapes via OMA, which gives the updated structural parameters from the simulated measurements, that is $\boldsymbol{\alpha}_0$ for the intact condition and $\boldsymbol{\alpha}$ for the damaged condition. Subtracting $\boldsymbol{\alpha}_0$ from $\boldsymbol{\alpha}$ gives the estimated damage severity at each possible damaged site, that is $\boldsymbol{\theta} = \boldsymbol{\alpha} - \boldsymbol{\alpha}_0$.

Table 6.1. Results of DMU and damage identification in step 1.

damage location	$\boldsymbol{\alpha}_0$	$\boldsymbol{\alpha}$	$\boldsymbol{\theta}$: estimated	$\boldsymbol{\theta}$: target
blade 1	0.07	-0.14	-0.21	-0.2
blade 2	0.00	-0.10	-0.10	-0.15
blade 3	0.05	-0.22	-0.27	-0.25
tower	0.03	-0.33	-0.36	-0.3

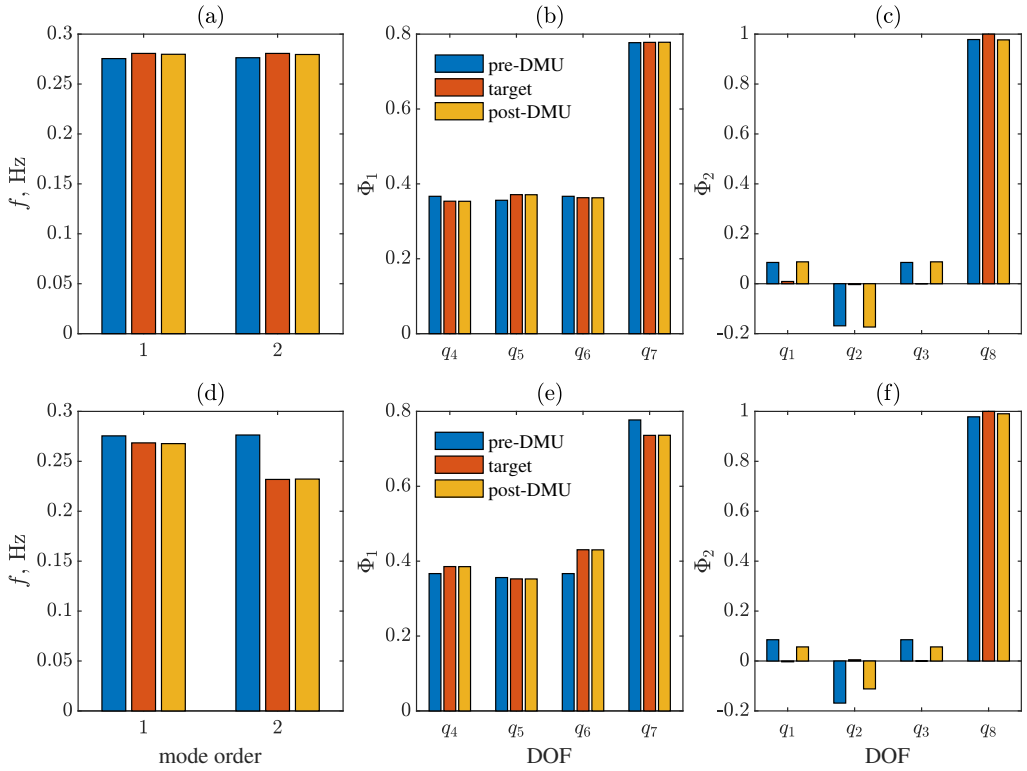


Figure 6.6. Comparison of modal properties from OMA and eigenanalysis using the generalized OWT model before and after DMU. OMA outputs serve as the targets of DMU. (a) to (c) are frequencies and mode shapes of the intact condition; (d) to (f) are frequencies and mode shapes of damage case 1.

Table 6.1 shows the results of DMU and damage identification of the above damaged case. It shows that, for the intact condition, DMU does not yield zero values for all locations, due to the limitation of OMA because of the wind and wave loads that are not strictly steady-state random excitations. However, this limitation will not significantly affect the performance of damage identification, as the damage severities are identified via

comparison. As shown in the 4^{th} column of Table 6.1, DMU accurately identifies the damage severities on all the OWT components, when compared with the target values in the 5^{th} column.

Figures 6.6 (a) to (f) compare the modal properties before DMU (pre-DMU), after DMU (post-DMU), and the target values. Figures 6.6 (a) to (c) represent the intact condition and Figures 6.6 (d) to (f) the mentioned damaged condition. The pre-DMU properties are calculated using the reference matrices; the post-DMU properties are calculated with the updated structural parameters $\boldsymbol{\alpha}$ or $\boldsymbol{\alpha}_0$; the target properties are obtained from OMA. Figure 6.6 (a) and (d) show that after DMU the calculated frequencies with the updated parameters are drawn closer to the target from the pre-DMU values. The difference is less obvious for the intact condition as shown in Figures 6.6 (a), as the pre-DMU properties are obtained from reference matrices of intact condition. The same phenomena happens to the mode shape as shown in the rest of subfigures. In summary, DMU tunes the structural parameters so that the estimated modal properties become more consistent with the measurements.

Table 6.1 lists the DMU and damage identification results for a general damage case with significant stiffness loss on each blade and the tower. However, in reality, damage usually initiates on a certain component (i.e., on a single blade), which may lead to catastrophic failure of the whole wind turbine if not detected timely. Therefore, the rest of this section considers three cases with structural damage occurring to certain but not all components. It should be noted that the damage condition in these three cases correspond to three damage cases in Section 6.4 for detailed local damage identification. In this way, the results of two steps can be connected and jointly demonstrate the effectiveness of the

proposed damage identification strategy for OWT. In detail, the first case (case 1: blade 1) corresponds to the case 1 in Table 6.3; the second case (case 1: tower) corresponds to the case 1 in Table 6.4; the third case (case 3: combination) corresponds to the case in Table 6.5. The details of these damage cases will be introduced in Section 6.4. The damage severity of each case in the analytical model is calculated by integrating the reduced structural stiffness along the corresponding mode shape curvatures [177].

Table 6.2 lists the results of damage identification for the three cases described above. It shows that DMU successfully identifies the damaged component(s) by yielding a considerable damage severity on the target damaged component(s) and ignorable stiffness change on intact components. Moreover, similar to the case in Table 6.1, DMU yields accurate damage severities to all three damage cases with a maximum relative error of 28.6% and a minimum of 0%. Further damage localization and quantification will be implemented in step 2 and the results are presented in Section 4.

Table 6.2. Results of damage identification in step 1: more cases.

damage location	case 1: blade 1		case 1: tower		case 3: combination	
	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated
blade 1	-0.18	-0.17	0.00	-0.02	-0.18	-0.18
blade 2	0.00	0.02	0.00	0.06	0.00	0.05
blade 3	0.00	-0.01	0.00	-0.03	0.00	-0.03
tower	0.00	0.00	-0.14	-0.10	-0.14	-0.10

6.4. Damage Localization and Quantification via Updating a Detailed FE Model

With damaged components identified in step 1 as presented in Section 3, this section further localizes and quantifies the damage on each of the damaged component(s). An FE model based on beam element is first established for the OWT using ANSYS 18.1. The following procedures are similar to that used in step 1, including numerical simulation, OMA, FE model updating via DMU, and damage localization and quantification. One challenge in step 2 is the complexity of the FE model that contains more than 1000 DOFs, which largely amplifies the difficulty of solving the optimization problem in DMU. To address this challenge, this study uses a substructural model of the identified damaged component, i.e., a blade or the tower for DMU in step 2. In the substructural model of a blade, fixed boundary condition is applied at the nacelle end; in the case of the tower, the masses of the three blades are added to the nacelle. The rest of this section will evaluate the the effect of this model simplification on damage identification result. Moreover, considering the independence between the in-plane and out-of-plane motions of the blade and the tower, the FE model of each substructure is further simplified as a 2D model. In this way, the DOFs of the model used for DMU in step 2 is limited within 100, which greatly reduces the number of dimension of the optimization problem.

6.4.1. FE Model Establishment and Numerical Simulations

Figure 6.7 shows the FE model of the OWT established in ANSYS mainly using beam elements. The effects of soil are simulated with COMBIN 14 elements with the real constants assigned as spring and damping coefficients. The nacelle and soil masses are simulated with MASS 21 elements. In total, the FE model has 172 elements, with 2 of them

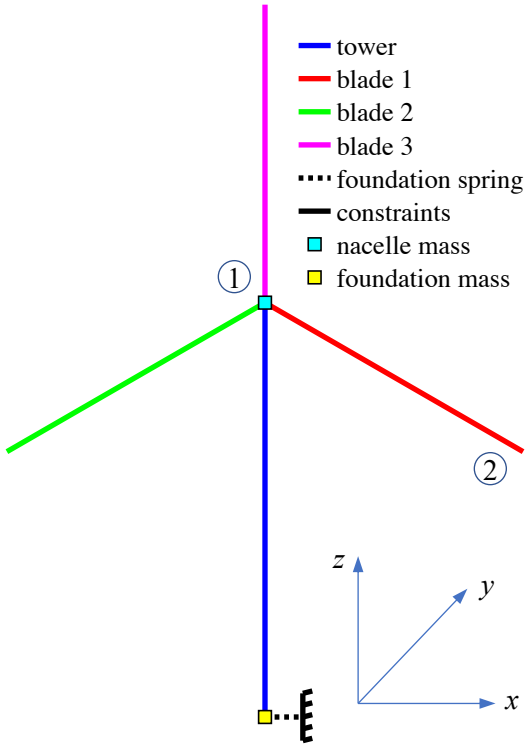


Figure 6.7. FE model of an OWT. The nacelle and the tip of blade 1 are denoted as locations 1 and 2 for future reference.

being MASS 21, 6 COMBIN 14, and the rest BEAM 4.

With the FE model established, the generated wind and wave loading in Section 3.1 are applied to the corresponding beam elements, as shown in Figure 6.8. Figure 6.9 shows some example responses of the intact OWT under certain excitations at the nacelle and blade tip. Simulated accelerations will be used for OMA in Section 4.2.

6.4.2. Operational Modal Analysis

Due to the complexity of the FE OWT model, it is challenging to extract the modal properties of the whole OWT from simulated responses at all measured DOFs. As pointed out in the proposed methodology in Section 2, the identified damaged component in step 1 can be a blade or the tower. Considering the relative independence between the blades' deformation and the tower's motion and the independence between the in-plane

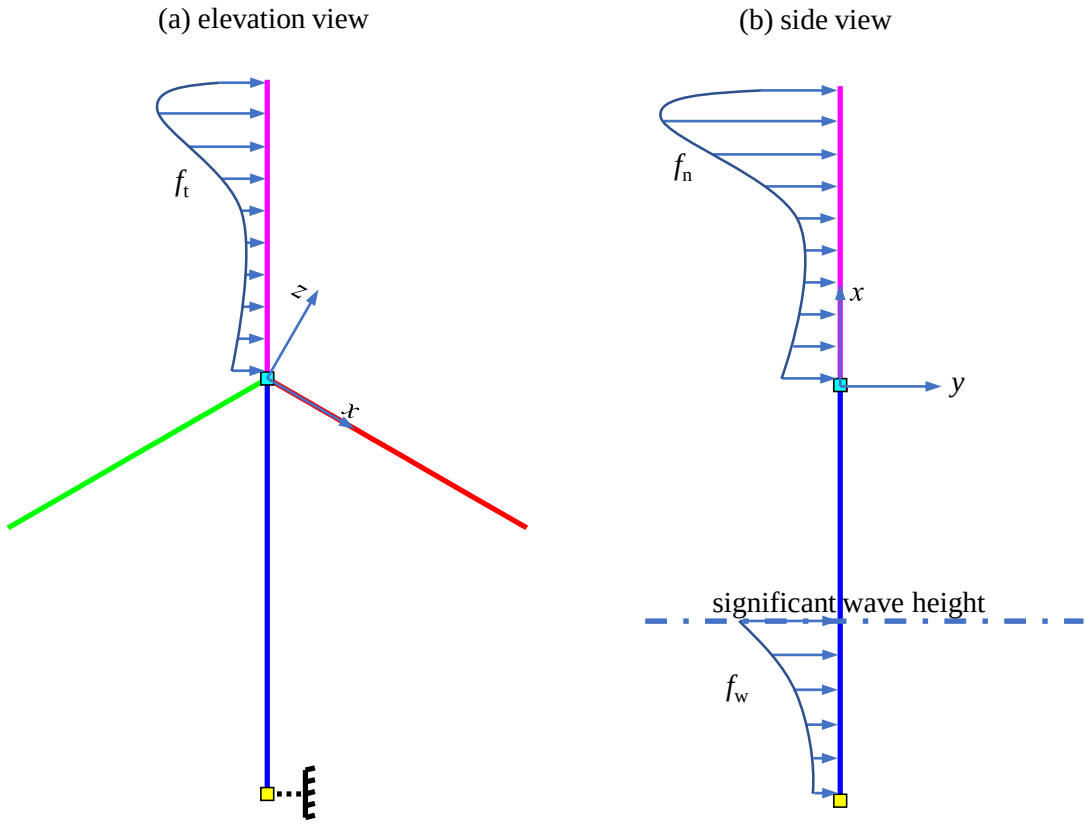


Figure 6.8. Wind and wave loading acing on the OWT. (a) the elevation view showing the in-plane loading acting on structural component(s); (b) the side view showing the out-of-plane loading acting on structural component(s). f_n denotes wind load component normal to the plane in the local y direction; f_t denotes wind load component tangential to the plane in the local z direction; f_w denotes wave load without misalignment in the local y direction. The local coordinate system is for each blade.

and out-of-plane motions, this study implements OMA separately for the tower and blade(s) and independently for in-plane and out-of-plane motions. Additionally, when damage happens to one blade, it may affect the relative motion between blades. Hence, it is beneficial to extract the modal properties from the responses of three blades and compare the results under different damage conditions. This section presents the modal properties of the three blades using eigen-analysis with respect to the FE OWT model and OMA of the structural response under intact and damaged conditions.

Figures 6.10 (a) to (f) compare the natural frequencies and mode shapes from FE

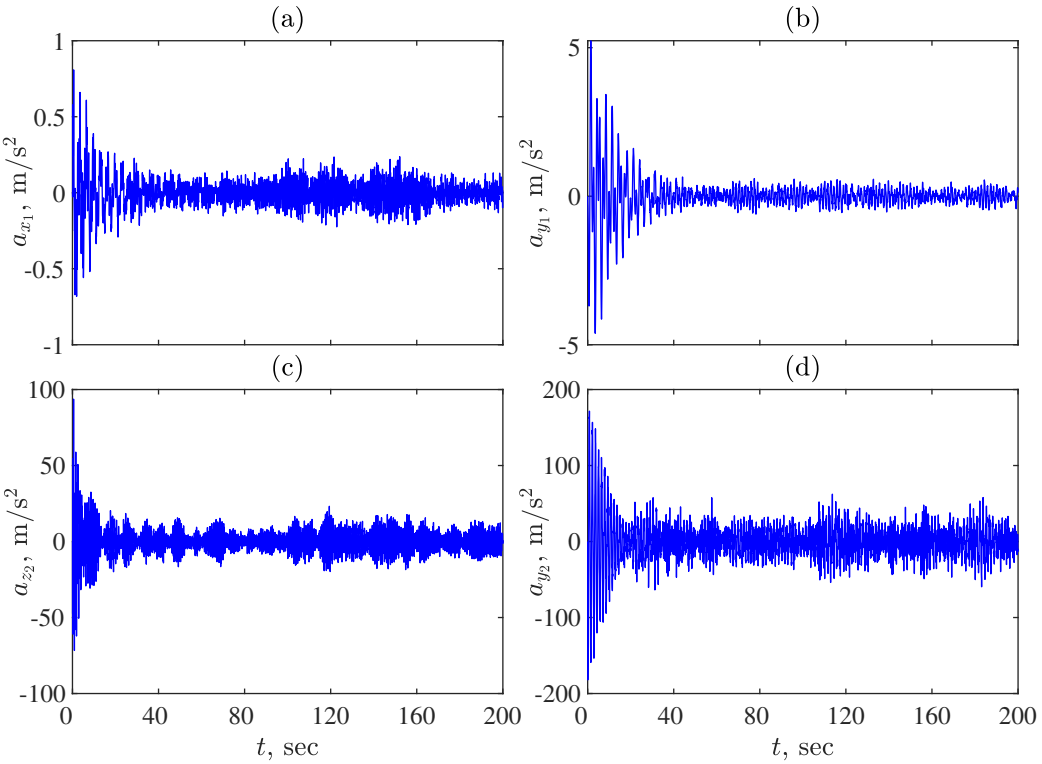


Figure 6.9. Structural responses of wind turbine under intact condition. (a) and (b) show accelerations of the nacelle in the global coordinates as shown in Figure 6.7; (c) and (d) are accelerations at the tip of blade 1 in local coordinates as shown in Figure 6.8.

analysis (FEA) and OMA of the intact and one damaged condition. One can see that the frequencies from FEA and OMA are very close for each mode with a maximum difference of 0.01 Hz. The mode shapes from OMA match that from FEA well in most cases ((a), (c), (d), and (f)); however, slight ((e)) or large ((b)) deviation can also be observed. The observed deviations manifest the challenge of OMA using responses of more than one OWT components. Therefore, for FE model updating for damage identification in Section 4.3, OMA is rerun using simulated in-plane/out-of-plane responses measured from a certain component.

Figures 6.11 (a) to (c) compare the OMA outputs for the intact and one damaged condition. It shows that structural damage on a certain blade causes significant frequency

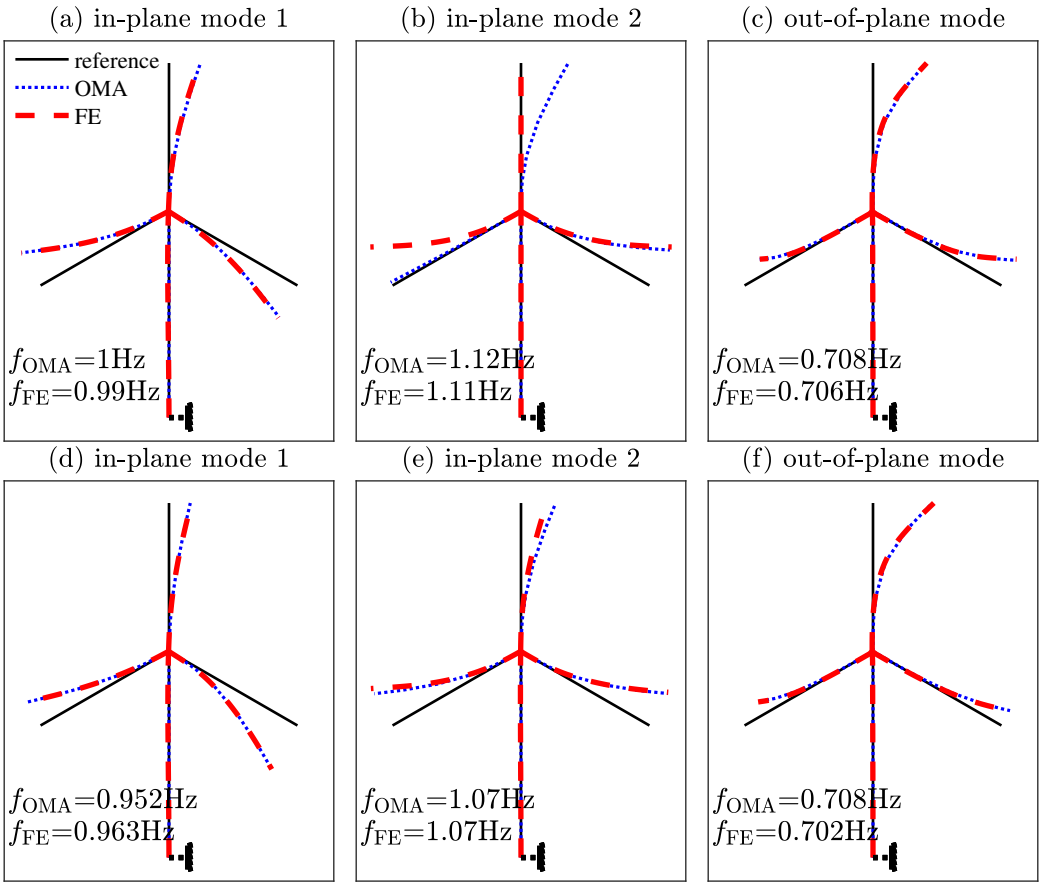


Figure 6.10. Comparison of mode shapes from OMA and FE model. (a)-(c): the intact condition; (d)-(f): the same damage condition with that in Figure 6.5. (a),(b),(d), and (e): the in-plane modes; (c) and (f): the out-of-plane modes plotted in the in-plane coordinate with their magnitudes. f_{OMA} is the natural frequency of a certain mode from OMA; f_{FE} is the natural frequency of that mode calculated from the FE model.

reduction and mode shape variation of the two in-plane modes ((a) and (b)). No significant change can be observed from the out-of-plane mode ((c)). It was noted that, despite the observed variations in modal properties, model updating is challenging using the whole FE OWT model.

6.4.3. FE Model Updating and Structural Damage Evaluation

This section implements damage identification using substructural FE models for a certain identified damaged component. To this end, OMA is first run using measured

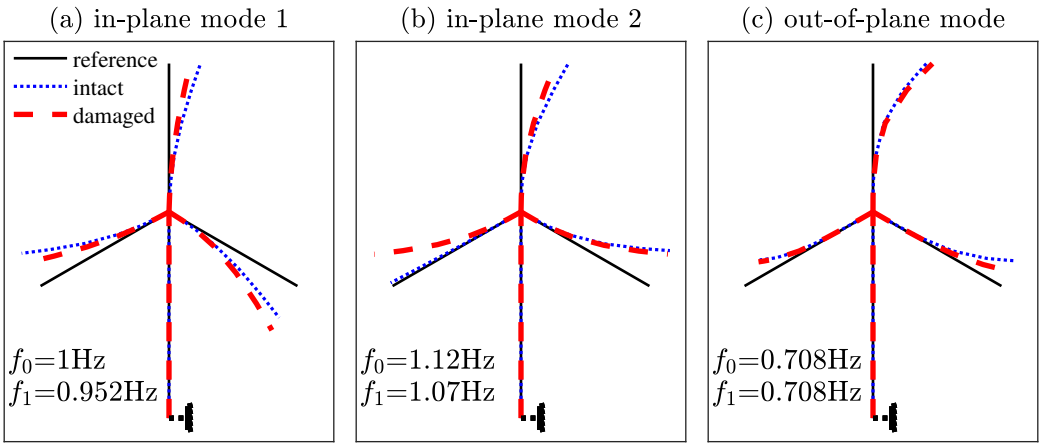


Figure 6.11. Comparison of mode shapes from OMA for intact and damaged conditions. The same damaged condition is used with that in Figure 6.5. (a) and (b): the in-plane modes; (c): the out-of-plane mode. f_0 is the natural frequency of the intact condition, and f_1 denotes that of the damaged condition.

responses from that component. Then FE model updating is performed via formulating the objective function and solving the optimization problem. Finally, structural damage is evaluated and the modal properties before and after model updating are compared. Three scenarios are considered in step 2 of the proposed methodology: (1) scenario 1: damage happens to blade 1; (2) scenario 2: damage happens to the tower; (3) scenario 3: damage happens to blade 1 and the tower. Three damage cases are simulated in scenarios 1 and 2, and one case in scenario 3. In step 2, out-of-plane modes for each component have better quality than in-plane modes in OMA and thus are used in DMU.

Scenario 1: Damage Happens to Blade 1

Figures 6.12 (a) to (f) compare the modal properties from OMA using the measured responses of blade 1 and that from FEA using the full FE OWT model and the substructural FE model of blade 1. It shows that the results of OMA match well with that from FEA with the full FE OWT model with a maximum frequency difference of 3%. This consistence indicates that OMA using responses of a blade yields good quality, especially

when compared with that using responses of three blades simultaneously (see Figure 6.10). However, FEA results using the substructural model are considerably different from that with the full FE model (with a 11.7% maximum frequency difference) and that of OMA, due to model simplification and reduction. This difference indicates that substructural modeling introduces modeling errors in model updating, which is one of the largest challenge in model-based SHM [54]. The rest of this section will investigate the adverse effects of modeling error using the results of DMU and damage evaluation.

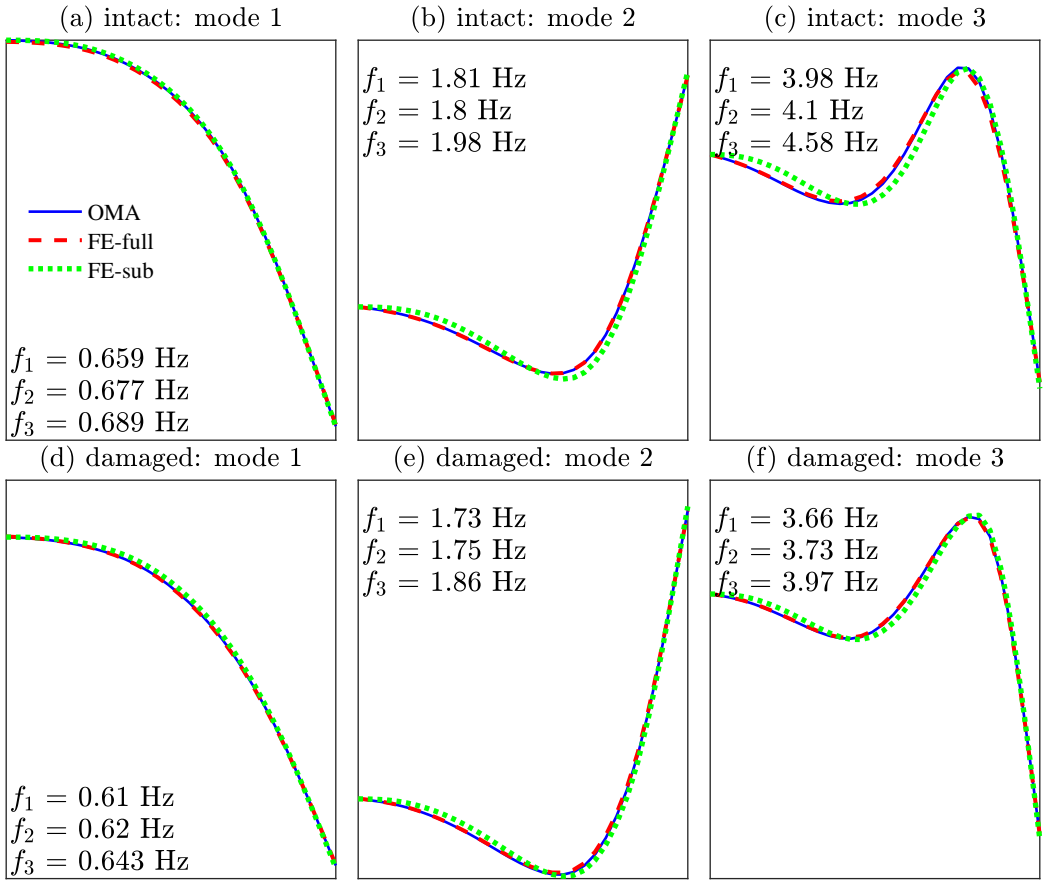


Figure 6.12. Comparison of modal properties from OMA and eigenanalysis using the full FE model and substructural FE model. Three out-of-plane modes are used for DMU and thus presented in this figure. (a) to (c): the intact condition; (d) to (f): case 1 with certain damage on blade 1. f_1, f_2, and f_3 are frequencies of a certain mode from OMA, eigenanalysis with the full FE model, and eigenanalysis with the substructural FE model, respectively.

Three cases with different damage locations on blade 1 are simulated in this section. Table 6.3 lists the results of damage identification via DMU and the target damage severities on each segment. It shows that DMU successfully localizes damage on the blade with significant evaluated damage severity on damaged segments except the one with 10% damage (S_5 in Case 2). Positive stiffness change are evaluated on intact segments that can be ignored in damage detection; hence, DMU does not yield false positive results in damage localization. Regarding the quantification of damage severity, it can be seen that the evaluated damage severities are considerably deviant from the target values because of the limitation of OMA quality and modeling errors. The quantification errors are especially large at S_1 (the root section of blade) due to the simplification of boundary conditions. The errors elsewhere range from 10% to 30%.

Table 6.3. Results of damage identification in scenario 1: damage happens to blade 1.

damage location	case 1		case 2		case 3	
	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated
S_1	-0.20	-0.05	0.00	0.03	-0.20	-0.08
S_2	0.00	0.01	-0.30	-0.24	-0.30	-0.26
S_3	-0.30	-0.21	0.00	0.04	-0.40	-0.30
S_4	0.00	0.04	-0.50	-0.39	0.00	0.04
S_5	0.00	0.04	-0.10	0.00	0.00	0.06
S_6	-0.50	-0.45	0.00	0.08	0.00	0.06

Figures 6.13 (a) to (f) compare the modal properties obtained via FEA with the pre-DMU random structural parameters, post-DMU updated parameter values, and the target quantities from OMA. It shows that frequencies and mode shapes obtained from DMU are closer to the target from the pre-DMU values, which confirms the effectiveness

of the proposed methodology of FE model updating in step 2. However, Figure 6.13 (c) shows an exception regarding the variation of modal properties after model updating: the frequency does not get closer to the target value and the mode shape is off from the target near the blade root. This exception can be attributed to modeling errors caused in sub-structural modeling, which is also consistent with the phenomena observed in Figure 6.12 (c).

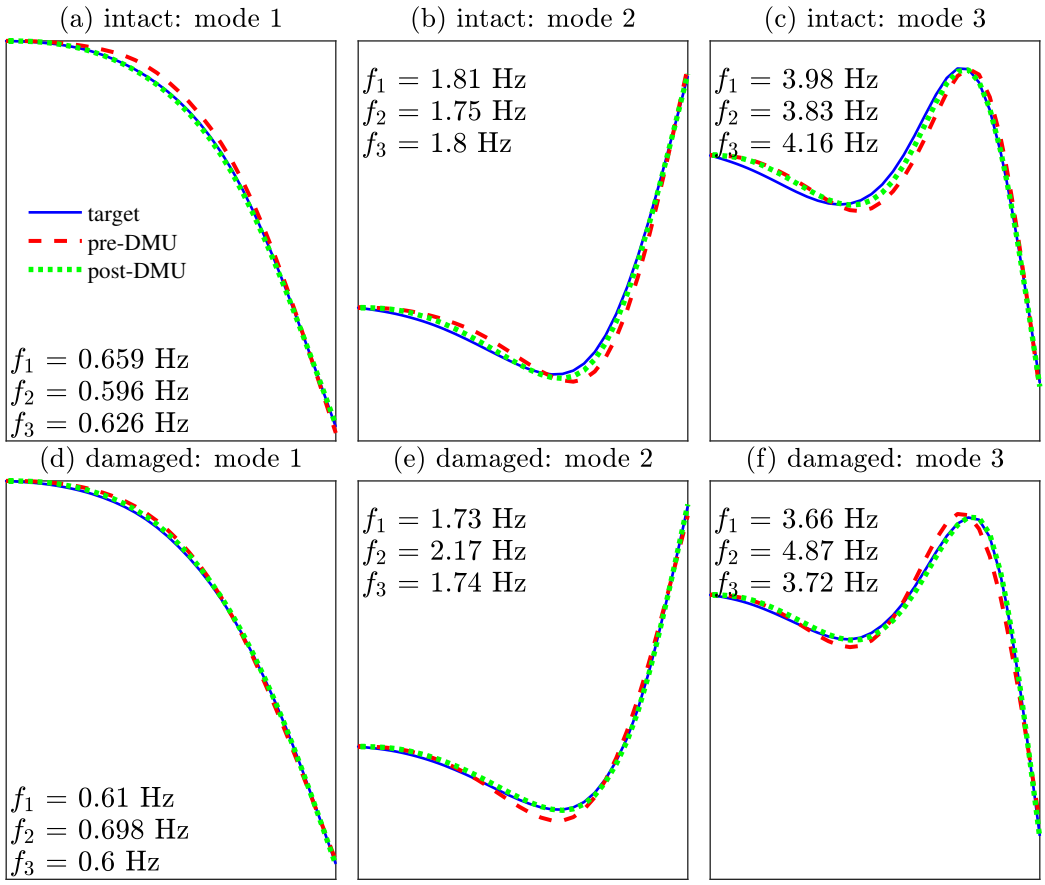


Figure 6.13. Comparison of modal properties from OMA and eigenanalysis using the sub-structural FE model before and after DMU. The outputs of OMA serve as the target of DMU. f_1 is the frequency of a certain mode from OMA; f_2, and f_3 are frequencies of the substructural FE model before and after DMU, respectively.

Scenario 2: Damage Happens to the Tower

Figures 6.14 (a) to (f) compare the modal properties obtained from FEA using the

full FE model and the substructural FE model of the tower and that from OMA. Similar to the case of blade as shown in Figure 6.12, FEA with full model yields mode shapes that are closer to that from OMA than that with substructural model in most cases. Nevertheless, frequencies from the substructural model are closer to OMA outputs in some cases as shown in Figure 6.14 (a), (b), (c), and (f).

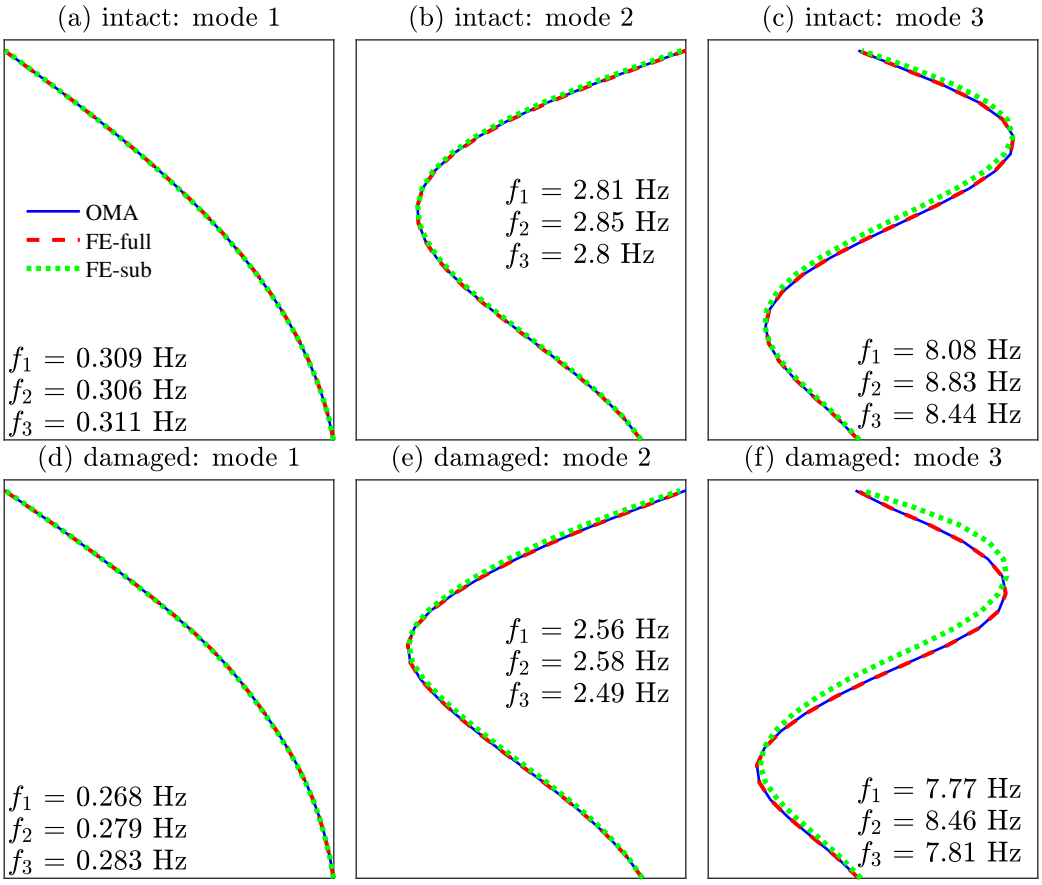


Figure 6.14. Comparison of modal properties from OMA and eigenanalysis using the full FE model and substructural FE model. Three out-of-plane modes are used for DMU and thus presented in this figure. (a) to (c): the intact condition; (d) to (f): case 1 with certain damage on the tower. f_1, f_2, and f_3 are frequencies of a certain mode from OMA, eigenanalysis with full FE model, and eigenanalysis with substructural FE model, respectively.

Table 6.4 lists the results of damage identification in scenario 2. It shows that except at S_6, DMU yields accurate damage quantities at all damaged locations including the

segment with only 10% damage, with the maximum relative error being 33.3% and the mean error 12.2%. However, when damage happens to S_6, the segment below the nacelle, DMU fails in damage identification. Again, this failure should be attributed to modeling error that results when establishing the substructural model of the tower. Lumping masses of blades at the nacelle ignores the influence of blades' structural response on the tower's response, and this influence is the most obvious for the top segment of the tower.

Table 6.4. Results of damage identification in scenario 2: damage happens to the tower.

damage location	case 1		case 2		case 3	
	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated
S_1	0.00	0.02	-0.30	-0.40	-0.10	-0.09
S_2	-0.30	-0.30	-0.50	-0.58	0.00	0.01
S_3	0.00	0.03	0.00	-0.09	0.00	0.02
S_4	-0.50	-0.50	-0.40	-0.48	-0.30	-0.29
S_5	0.00	0.12	0.00	0.06	-0.20	-0.17
S_6	-0.40	0.15	0.00	0.55	0.00	0.46

Figures 6.15 (a) to (f) compare the pre-DMU and post-DMU modal properties with target from OMA. Generally, frequencies and mode shape approach the target as a result of DMU. Differing from the case of blade, post-DMU mode shapes do not exhibit obvious deviation from the target, especially at the top segment, where DMU yields the largest error in damage quantity.

To further investigate the cause of the error in DMU at the top segment of the tower, this section further analyzes the mode shape curvatures from OMA, FEA with different FE models, and FEA with pre-DMU and post-DMU structural parameter values. Intact condition is used in this analysis. Figures 6.16 (a) to (c) compare mode shape cur-

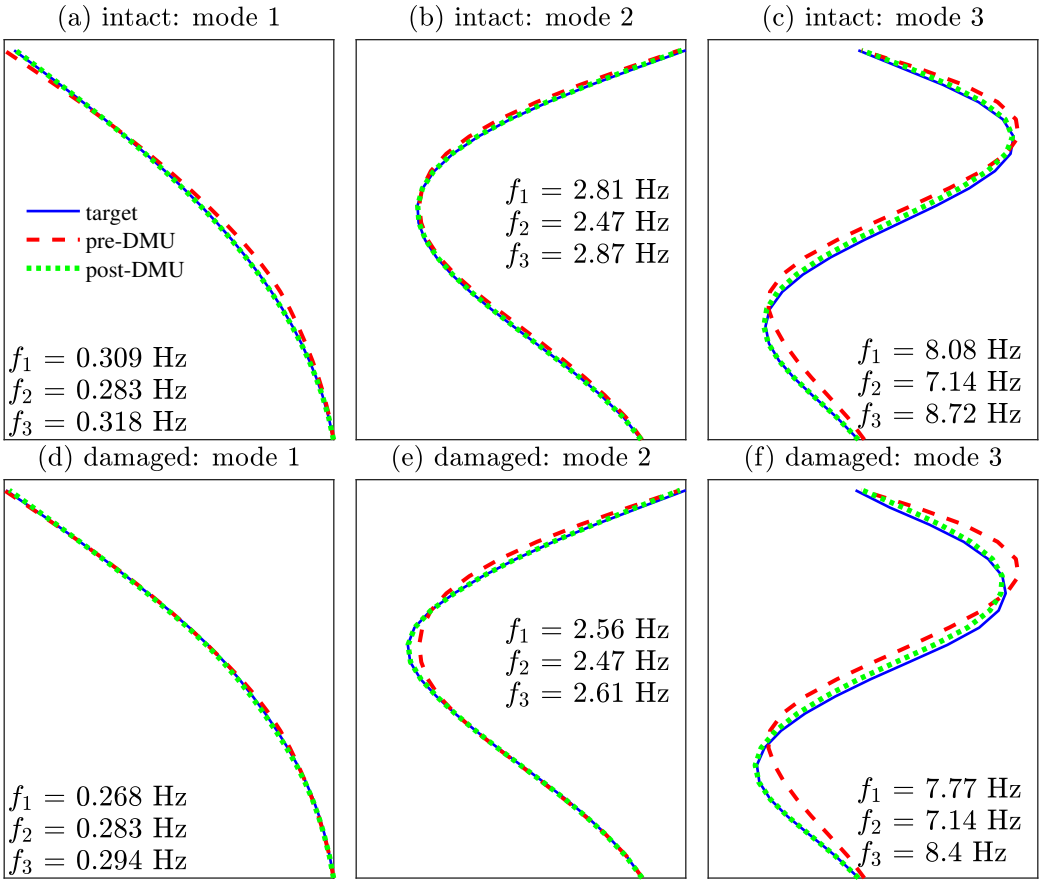


Figure 6.15. Comparison of modal properties from OMA and eigenanalysis using the substructural FE model before and after DMU. The outputs of OMA serve as the target of DMU. f_1 is the frequency of a certain mode from OMA; f_2, and f_3 are frequencies of the substructural FE model before and after DMU, respectively.

vatures from OMA and FEA with full and substructural models. It shows that the FEA output with the substructural model has slight ((a) and (b)) or considerable ((c)) deviation from that of OMA especially at the top segment when compared with that from FEA with the full model. The deviation is much larger than that of mode shapes as shown in Figures 6.14 (a) to (c). This difference indicates that error in substructural modeling has the largest adverse effects on mode shape curvatures. Figures 6.16 (d) to (f) compare the mode shape curvatures from OMA and that from pre-DMU and post-DMU FE models. It shows that DMU does not draw mode shape curvatures obviously closer to the target ones.

In mode 1 as shown in Figure 6.16 (d), it becomes even farther from the target at the top segment after DMU. Hence, the major factor causing significant error in damage identification at the top segment of tower is the insufficient modeling in the substructural tower model that leads to irreducible deviation of mode shape curvatures in model updating.

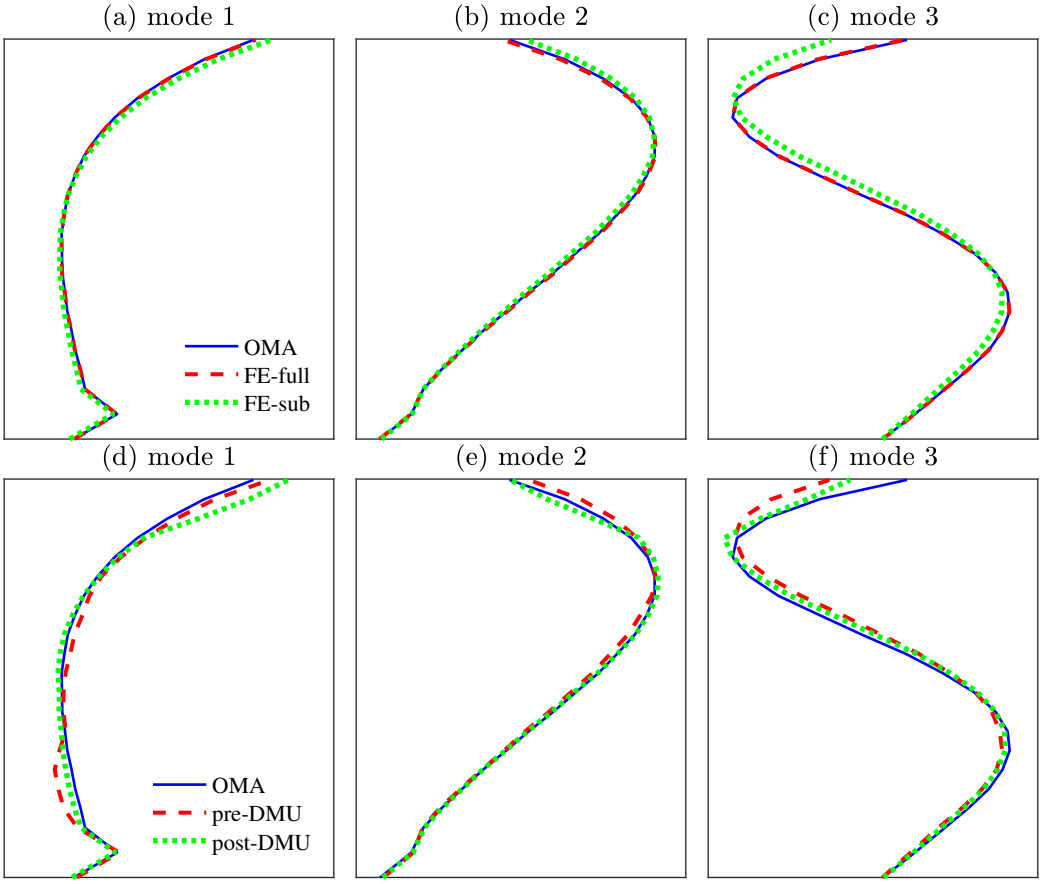


Figure 6.16. Comparison of mode shape curvatures from OMA and eigenanalysis using full and substructural FE models before and after DMU.

Scenario 3: Damage Happens to Blade 1 and the Tower

To demonstrate the effectiveness of the proposed methodology in identifying damages occurring simultaneously to different components, this section simulates a scenario with damage on certain segments of blade 1 and the tower respectively. Damage conditions of case 1 in scenario 1 and case 3 in scenario 2 are combined to formulate the dam-

age condition in this scenario. The results of damage identification are shown in Table 6.5. It shows that all the structural damage on segments of blade and tower can be accurately identified with considerable stiffness reduction evaluated on damaged segments and ignorable stiffness variation of intact segments. The performance of damage quantification is close to that when damage happens to one component as presented in Table 6.3 for blade and Table 6.4 for tower. The result presented in Table 6.5 indicates that when damage happens to segments of more than one component on OWT, damage evaluation can be implemented independently on each identified damaged component without weakening the performance.

Table 6.5. Results of damage identification in scenario 3: damage happens to blade 1 and the tower.

damage location	blade		tower	
	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated	$\boldsymbol{\theta}$: target	$\boldsymbol{\theta}$: estimated
S_1	-0.20	-0.07	-0.10	-0.09
S_2	0.00	0.00	0.00	0.01
S_3	-0.30	-0.22	0.00	0.02
S_4	0.00	0.04	-0.30	-0.29
S_5	0.00	0.04	-0.20	-0.17
S_6	-0.50	-0.44	0.00	0.38

6.5. Summary

This study proposes a two-step strategy of identifying structural damage on OWTs via FE model updating. Step 1 aims to detect the damaged component(s) that can be a certain blade(s) or the tower, without evaluating the exact damage location or quantity on the component(s). To this end, an analytical model of the OWT is established using the

Euler-Lagrangian equation. The analytical model has 12 DOFs considering only the first mode of each component. Wind and wave loads are generated and applied to the model to simulate the dynamic responses of the OWT. Subsequently, OMA (operational modal analysis) is implemented using the simulated responses, yielding modal properties of the model under intact and damaged conditions. Following OMA, model updating minimizes the objective function formulated with outputs of OMA and yields the updated structural parameters for each damage condition. Finally, the damaged component(s) can be determined via comparing the parameter values of a certain unknown condition with that of the intact model. The results indicate that the damaged component(s) can be successfully identified. Additionally, the modal properties calculated with updated structural parameters get closer to the OMA results from the initially assumed value.

With damaged component(s) detected in step 1, step 2 focues on accurate identification of the damage location and severity. For this purpose, an FE model of the OWT is built in ANSYS to simulate the structural response under given wind and wave loading. Moreover, a 2D substructural FE model is built for each component, which is finally used in model updating to reduce the dimension of the optimization problem and thus improve the efficiency of problem solving. The results of model updating show that when damage happens to segments of a certain blade, significant stiffness loss ($\geq$ 20%) can be successfully localized. Damage quantity is accurately evaluated with an accuracy of 70% to 90%, except at the root section where the error is larger than 50% due to critical modeling error introduced during substructural modeling. When damage occurs to the tower, modeling error causes severe challenge in damage evaluation at its top segment, where the mode shape curvatures do not agree with each other after model updating. However, the

performance elsewhere is not affected significantly with all structural damage correctly localized and damage severity accurately quantified with an average accuracy of 88.8%. When damage simultaneously happens to one blade and the tower, the damage condition on each component can be evaluated separately without considerable loss of accuracy. Numerical analysis in this study indicates that modeling error is the major challenge of damage identification on OWT via model updating. Modeling error is an inherent challenge in model updating, especially for complex structures like an OWT. Its adverse effects can be mitigated by complementing the present methodology with some direct monitoring techniques such as acoustic emission and ultrasonic testing at critical locations of the OWT. Additionally, the adverse influence of modeling error may be mitigated if the process of FE model updating can be guided in a way with statistical pattern recognition. The feasibility and efficiency of this idea will be investigated in future studies.

Chapter 7. Conclusions and Future Research

In this dissertation, two critical issues in SHM methods are investigated, including the modeling uncertainty in model-based methods and data insufficiency in data-driven methods. To this end, the physics-guided machine learning (PGML) and transfer learning guided Bayesian model updating (BMU-TL) methods are proposed to bridge the gap between model-based and data-driven SHM methods, aiming to integrate their merits and minimize the adverse influence of their shortcomings. In addition, multi-site damage identification in data-driven SHM is studied in this dissertation to overcome the challenge of the dearth of data. Multi-label classification (MLC) and constrained independent component analysis (cICA) methods are investigated to resolve the correlations among damage cases sharing damaged sites and thus can improve the damage identification accuracy. Finally, as a case study, a two-step strategy is proposed for damage identification of offshore wind turbines via FE model updating. This chapter presents a summary of the findings of this dissertation and a list of recommendations for future research in this field.

7.1. Conclusions

7.1.1. Model-based SHM

Model-based SHM via FE model updating is susceptible to modeling uncertainty/error due to model simplification and omission. Consequently, the results of structural identification, damage evaluation, and subsequent response prediction are all biased from the truth. The case study on offshore wind turbine shows the influence of modeling error in substructural modeling on the performance of model updating and damage evalua-

tion. In this dissertation, two methods are proposed to mitigate the influence of modeling error on the performance of FE model updating: 1) PGML; 2) BMU-TL. In PGML, the results of FE model updating are incorporated into a neural network (NN) model to have a physics-guided neural network (PGNN) model, with an established FE model and its updated outputs regarded as an implicit representation of underlying physics of the target structure. It is indicated that modifications of the NN model with physics guidance can improve physics-related consistency of the learned model. Moreover, with the most probable damage location identified from PGML, FE model updating can be rerun with constrained targets, which can considerably improve the efficiency of solving the optimization problem and increase its accuracy. Hence, the adverse effects of modeling error can be significantly mitigated. On the other hand, the domain adaptation based transfer learning technique is used to guide the direction of structural parameter optimization in Bayesian model updating. Domain adaptation is realized with the adaptation-regularization based transfer learning method that minimizes the influence of domain difference in pattern recognition through regularizing the distribution and manifold discrepancy. The results from both numerical and experimental case studies show that transfer learning can help identify the correct damage locations and thus direct and regularize the model updating process, so that the influence of modeling error can be minimized.

7.1.2. Data-driven SHM

The major challenge in data-driven SHM lies in data insufficiency with respect to both data quantity and quality, which leads to the overfitting problem in pattern recognition. This dissertation research attempts to solve the data insufficiency problem from two different perspectives: 1) PGML introduces physics guidance from FE model updating to

improve the physics-related consistency of the pattern recognition results which would otherwise be overfitted; 2) Transfer learning, as a means of domain adaptation, borrows data and knowledge learned from numerical modeling and simulations for damage evaluation on experimental structures, considering the discrepancy between domains. Both these two approaches have been numerically and experimentally proved effective in improving the performance of data-driven SHM with insufficient measured data. Moreover, both PGML and TL-BMU provide a scheme of integrating model-based and data-driven SHM for better performance of damage diagnosis.

Multi-site damage identification has significant importance in structural health monitoring. The main challenge of data-driven multi-site damage detection lies in the ambiguity caused by information mixture when damage happens at multiple locations. To address this challenge, this dissertation proposes using MLC and cICA methods in the framework of data-driven SHM. Both these two methods efficiently decompose the inter-correlation between damage cases and thus improve the performance of multi-site damage identification. Moreover, both of them demonstrate the potential of relieving the burden of data-driven SHM from data incompleteness/insufficiency for pattern recognition.

7.2. Future Research

Based on the conclusions drawn in this dissertation, ideas for future research are presented as follows. They are proposed to extend the work that has been done in this dissertation.

1. In PGML, the results of FE model updating is regarded as an implicit representation

of underlying physics. However, it is noted by the author that FE model updating outputs may not be the best representation of scientific rules for structural damage evaluation, which will be further investigated in the future research.

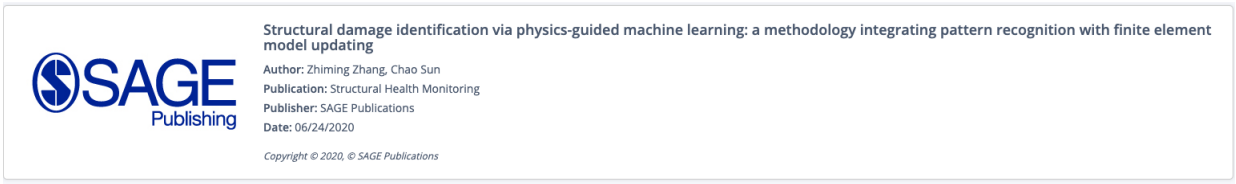
2. In PGML and BMU-TL, only single-site damage cases are considered. Multi-site damage identification requires incorporating corresponding machine learning methods, such as MLC, into the frameworks of PGML and BMU-TL. This work will also be taken into account in future studies.

3. The case study on wind turbine monitoring shows that modeling error is the major challenge of damage identification via model updating. In future studies, PGML ad BMU-TL methods will be used to solve this issue and improve the damage identification performance on wind turbines. Additionally, when experimental data is available in the future, the two-step methodology will be validated using experimental data.

4. A numerical study on multi-site damage identification with cICA method is presented in this dissertation. Future work will validate the proposed methodology using experimental and field testing data on real structures via applying multiple types of loading, including harmonic excitations, impact loading, and environmental excitations, to prove the practical feasibility and generality of the proposed methodology.

5. Global features extracted from modal properties are used in this thesis for structural damage identification, which might be limited in detecting small-scale defects that do not cause remarkable stiffness reduction, such as fatigue cracking. Future work will consider developing and incorporating good features sensitive to this type of damage.

Appendix A. Permissions



Structural damage identification via physics-guided machine learning: a methodology integrating pattern recognition with finite element model updating

Author: Zhiming Zhang, Chao Sun
Publication: Structural Health Monitoring
Publisher: SAGE Publications
Date: 06/24/2020

Copyright © 2020, © SAGE Publications

If you are a SAGE journal author requesting permission to reuse material from your journal article, please note you may be able to reuse your content without requiring permission from SAGE. Please review SAGE's author re-use and archiving policies at https://us.sagepub.com/en-us/nam/journal-author-archiving-policies-and-re-use for more information.

If your request does not fall within SAGE's reuse guidelines, please proceed with submitting your request by selecting one of the other reuse categories that describes your use. Please note, a fee may be charged for reuse of content requiring permission. Please contact permissions@sagepub.co.uk if you have questions.

BACK CLOSE WINDOW

Multi-site structural damage identification using a multi-label classification scheme of machine learning

Author: Zhiming Zhang,Chao Sun
Publication: Measurement
Publisher: Elsevier
Date: 15 March 2020

© 2020 Elsevier Ltd. All rights reserved.

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

BACK CLOSE WINDOW

Bibliography

[1] D. Balageas, C.-P. Fritzen, A. Güemes, Structural health monitoring, Vol. 90, John Wiley & Sons, 2010.

[2] A. Teughels, J. Maeck, G. De Roeck, Damage assessment by fe model updating using damage functions, Computers & structures 80 (25) (2002) 1869–1879.

[3] M. Song, I. Behmanesh, B. Moaveni, C. Papadimitriou, Modeling error estimation and response prediction of a 10-story building model through a hierarchical bayesian model updating framework, Frontiers in Built Environment 5 (2019) 7.

[4] C. R. Farrar, K. Worden, Structural health monitoring: a machine learning perspective, John Wiley & Sons, 2012.

[5] R. Barthorpe, G. Manson, K. Worden, On multi-site damage identification using single-site training data, Journal of Sound and Vibration 409 (2017) 43–64.

[6] T. Stolarski, Y. Nakasone, S. Yoshimoto, Engineering analysis with ANSYS software, Butterworth-Heinemann, 2018.

[7] C. R. Farrar, K. Worden, Structural health monitoring: a machine learning perspective, John Wiley & Sons, 2012.

[8] Z. Zhang, C. Sun, Multi-site structural damage identification using a multi-label classification scheme of machine learning, Measurement (2020) 107473.

[9] R. J. Barthorpe, On model-and data-based approaches to structural health monitoring, 2010, University of Sheffield.

[10] Z. Zhang, C. Sun, A numerical study on multi-site damage identification: a data-driven method via constrained independent component analysis, Structural Control and Health Monitoring (In press).

[11] Y. Ying, J. H. Garrett Jr, I. J. Oppenheim, L. Soibelman, J. B. Harley, J. Shi, Y. Jin, Toward data-driven structural health monitoring: application of machine learning and signal processing to damage detection, Journal of Computing in Civil Engineering 27 (6) (2012) 667–680.

[12] K. Worden, G. Manson, The application of machine learning to structural health monitoring, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 365 (1851) (2007) 515–537.

[13] M. Sanayei, B. Arya, E. M. Santini, S. Wadia-Fascetti, Significance of modeling error in structural parameter estimation, Computer-Aided Civil and Infrastructure Engi-

neering 16 (1) (2001) 12–27.

[14] M. I. Friswell, Damage identification using inverse methods, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365 (1851) (2006) 393–410.

[15] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou, Hierarchical bayesian model updating for structural identification, Mechanical Systems and Signal Processing 64 (2015) 360–376.

[16] A. Karpatne, G. Atluri, J. H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, V. Kumar, Theory-guided data science: A new paradigm for scientific discovery from data, IEEE Transactions on Knowledge and Data Engineering 29 (10) (2017) 2318–2331.

[17] J. Kawale, S. Liess, A. Kumar, M. Steinbach, P. Snyder, V. Kumar, A. R. Ganguly, N. F. Samatova, F. Semazzi, A graph-based approach to find teleconnections in climate data, Statistical Analysis and Data Mining: The ASA Data Science Journal 6 (3) (2013) 158–179.

[18] J.-X. Wang, J.-L. Wu, H. Xiao, Physics-informed machine learning for predictive turbulence modeling: Using data to improve rans modeled reynolds stresses, arXiv preprint arXiv:1606.07987 (2016) 1041–4347.

[19] G. Hautier, C. C. Fischer, A. Jain, T. Mueller, G. Ceder, Finding nature's missing ternary oxide compounds using machine learning and density functional theory, Chemistry of Materials 22 (12) (2010) 3762–3767.

[20] L. Li, J. C. Snyder, I. M. Pelaschier, J. Huang, U.-N. Niranjan, P. Duncan, M. Rupp, K.-R. Müller, K. Burke, Understanding machine-learned density functionals, International Journal of Quantum Chemistry 116 (11) (2016) 819–833.

[21] A. Karpatne, W. Watkins, J. Read, V. Kumar, Physics-guided neural networks (pgnn): An application in lake temperature modeling, arXiv preprint arXiv:1710.11431.

[22] Y. Wang, X. Dong, D. Li, Y. Otsuki, SMU: MATLAB package for structural model updating, version 1.1, `https://github.com/ywang-structures/Structural-Model-Updating` (2019).

[23] E. Figueiredo, G. Park, J. Figueiras, C. Farrar, K. Worden, Structural health monitoring algorithm comparisons using standard data sets, Tech. rep., Los Alamos National Laboratory (LANL), Los Alamos, NM (United States) (2009).

[24] X. Dong, Y. Wang, Formulation and optimization algorithm comparison for the FE

model updating of large-scale models., https://github.com/ywang-structures/Structural-Model-Updating (2018).

[25] P. Kaminski, The approximate location of damage through the analysis of natural frequencies with artificial neural networks, Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering 209 (2) (1995) 117–123.

[26] R. L. Wilby, T. Wigley, D. Conway, P. Jones, B. Hewitson, J. Main, D. Wilks, Statistical downscaling of general circulation model output: A comparison of methods, Water resources research 34 (11) (1998) 2995–3008.

[27] C. M. Bishop, Pattern recognition and machine learning, springer, 2006.

[28] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (7553) (2015) 436–444.

[29] N. Ketkar, Introduction to PyTorch, in: Deep learning with python, Springer, 2017, pp. 195–208.

[30] P. Ramachandran, B. Zoph, Q. V. Le, Searching for activation functions, arXiv preprint arXiv:1710.05941.

[31] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.

[32] Y. Yao, L. Rosasco, A. Caponnetto, On early stopping in gradient descent learning, Constructive Approximation 26 (2) (2007) 289–315.

[33] S. Patro, K. K. Sahu, Normalization: A preprocessing stage, arXiv preprint arXiv:1503.06462.

[34] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, Journal of machine learning research 13 (Feb) (2012) 281–305.

[35] H. Sun, R. Betti, A hybrid optimization algorithm with bayesian inference for probabilistic model updating, Computer-Aided Civil and Infrastructure Engineering 30 (8) (2015) 602–619.

[36] R. Brincker, L. Zhang, P. Andersen, Modal identification of output-only systems using frequency domain decomposition, Smart materials and structures 10 (3) (2001) 441.

[37] M. Friswell, J. E. Mottershead, Finite element model updating in structural dynamics, Vol. 38, Springer Science & Business Media, 2013.

[38] A. Teughels, G. De Roeck, Structural damage identification of the highway bridge z24 by fe model updating, Journal of Sound and Vibration 278 (3) (2004) 589–610.

[39] R. Astroza, A. Alessandri, Effects of model uncertainty in nonlinear structural finite element model updating by numerical simulation of building structures, Structural Control and Health Monitoring 26 (3) (2019) e2297.

[40] J. L. Beck, Bayesian system identification based on probability logic, Structural Control and Health Monitoring 17 (7) (2010) 825–847.

[41] M. Song, B. Moaveni, C. Papadimitriou, A. Stavridis, Accounting for amplitude of excitation in model updating through a hierarchical bayesian approach: Application to a two-story reinforced concrete building, Mechanical Systems and Signal Processing 123 (2019) 68–83.

[42] R. Astroza, A. Alessandri, J. P. Conte, A dual adaptive filtering approach for nonlinear finite element model updating accounting for modeling uncertainty, Mechanical Systems and Signal Processing 115 (2019) 782–800.

[43] J.-A. Goulet, I. F. Smith, Structural identification with systematic errors and unknown uncertainty dependencies, Computers & structures 128 (2013) 251–258.

[44] B. Moaveni, I. Behmanesh, Effects of changing ambient temperature on finite element model updating of the dowling hall footbridge, Engineering Structures 43 (2012) 58–68.

[45] I. Behmanesh, B. Moaveni, Accounting for environmental variability, modeling errors, and parameter estimation uncertainties in structural identification, Journal of Sound and Vibration 374 (2016) 92–110.

[46] A. Nozari, I. Behmanesh, S. Yousefianmoghadam, B. Moaveni, A. Stavridis, Effects of variability in ambient vibration data on model updating and damage identification of a 10-story building, Engineering Structures 151 (2017) 540–553.

[47] H. Sohn, K. H. Law, A bayesian probabilistic approach for structure damage detection, Earthquake engineering & structural dynamics 26 (12) (1997) 1259–1281.

[48] M. Sanayei, A. Khaloo, M. Gul, F. N. Catbas, Automated finite element model updating of a scale bridge model using measured static and modal test data, Engineering Structures 102 (2015) 66–79.

[49] I. Behmanesh, B. Moaveni, C. Papadimitriou, Probabilistic damage identification of a designed 9-story building using modal data in the presence of modeling errors, Engineering Structures 131 (2017) 542–552.

[50] E. L. Droguett, A. Mosleh, Bayesian methodology for model uncertainty using model performance data, Risk Analysis: An International Journal 28 (5) (2008) 1457–1476.

[51] T. D. Brown, Finite element modeling in musculoskeletal biomechanics, Journal of applied biomechanics 20 (4) (2004) 336–366.

[52] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: A review, Mechanical Systems and Signal Processing 56 (2015) 123–149.

[53] E. Simoen, B. Moaveni, J. P. Conte, G. Lombaert, Uncertainty quantification in the assessment of progressive damage in a 7-story full-scale building slice, Journal of Engineering Mechanics 139 (12) (2013) 1818–1830.

[54] J. L. Beck, L. S. Katafygiotis, Updating models and their uncertainties. i: Bayesian statistical framework, Journal of Engineering Mechanics 124 (4) (1998) 455–461.

[55] J. Brynjarsdottir, A. OHagan, Learning about physical parameters: The importance of model discrepancy, Inverse Problems 30 (11) (2014) 114007.

[56] B. Titurus, M. I. Friswell, L. Starek, Damage detection using generic elements: Part ii. damage detection, Computers & structures 81 (24-25) (2003) 2287–2299.

[57] M. W. Vanik, J. L. Beck, S. Au, Bayesian probabilistic approach to structural health monitoring, Journal of Engineering Mechanics 126 (7) (2000) 738–745.

[58] J.-A. Goulet, I. F. Smith, Probabilistic model falsification for structural identification, in: Proceedings of 11th International Conference on Structural Safety & Reliability (ICOSSAR), New York, USA, 2013, no. CONF, 2013.

[59] B. Goller, G. Schueller, Investigation of model uncertainties in bayesian structural model updating, Journal of sound and vibration 330 (25) (2011) 6122–6136.

[60] L. S. Katafygiotis, J. L. Beck, Updating models and their uncertainties. ii: Model identifiability, Journal of Engineering Mechanics 124 (4) (1998) 463–467.

[61] I. Behmanesh, S. Yousefianmoghadam, A. Nozari, B. Moaveni, A. Stavridis, Uncertainty quantification and propagation in dynamic models using ambient vibration measurements, application to a 10-story building, Mechanical Systems and Signal Processing 107 (2018) 502–514.

[62] K. Alvin, W. Oberkampf, K. Diegert, B. Rutherford, Uncertainty quantification in computational structural dynamics: a new paradigm for model validation, in: Society for Experimental Mechanics, Inc, 16 th International Modal Analysis Conference., Vol. 2, 1998, pp. 1191–1198.

[63] R. Zhang, S. Mahadevan, Model uncertainty and bayesian updating in reliability-based inspection, Structural Safety 22 (2) (2000) 145–160.

[64] F. Zouaoui, J. R. Wilson, Accounting for input model and parameter uncertainty in simulation, in: Proceedings of the 33nd conference on Winter simulation, IEEE Computer Society, 2001, pp. 290–299.

[65] E. Zio, G. Apostolakis, Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories, Reliability Engineering & System Safety 54 (2-3) (1996) 225–241.

[66] J. M. Reinert, G. E. Apostolakis, Including model uncertainty in risk-informed decision making, Annals of nuclear energy 33 (4) (2006) 354–369.

[67] D. Anderson, K. Burnham, Model selection and multi-model inference, Second. NY: Springer-Verlag (2004) 63.

[68] J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, J. Wortman, Learning bounds for domain adaptation, in: Advances in neural information processing systems, 2008, pp. 129–136.

[69] S. J. Pan, Q. Yang, A survey on transfer learning, IEEE Transactions on knowledge and data engineering 22 (10) (2009) 1345–1359.

[70] S. J. Pan, I. W. Tsang, J. T. Kwok, Q. Yang, Domain adaptation via transfer component analysis, IEEE Transactions on Neural Networks 22 (2) (2010) 199–210.

[71] X. Liu, K. Worden, On the application of domain adaptation in SHM, in: Special Topics in Structural Dynamics and Experimental Techniques, Volume 5. Conference Proceedings of the Society for Experimental Mechanics Series., Vol. 5, 2020, pp. 111–122.

[72] S. Sehgal, H. Kumar, Structural dynamic model updating techniques: A state of the art review, Archives of computational methods in engineering 23 (3) (2016) 515–533.

[73] Y. Bao, Z. Chen, S. Wei, Y. Xu, Z. Tang, H. Li, The state of the art of data science and engineering in structural health monitoring, Engineering 5 (2) (2019) 234–242.

[74] Z. Zhang, C. Sun, Structural damage identification via physics-guided machine learning: A methodology integrating pattern recognition with finite element model updating, Structural Health Monitoring (In press).

[75] W. R. Gilks, S. Richardson, D. Spiegelhalter, Markov chain Monte Carlo in practice, Chapman and Hall/CRC, 1995.

[76] M. Long, J. Wang, G. Ding, S. J. Pan, S. Y. Philip, Adaptation regularization: A general framework for transfer learning, IEEE Transactions on Knowledge and Data Engineering 26 (5) (2013) 1076–1089.

[77] Z. Zhang, C. Sun, C. Li, M. Sun, Vibration based bridge scour evaluation: A data-driven method using support vector machines, Structural Monitoring and Maintenance 6 (2) (2019) 125–145.

[78] S. W. Doebling, F. M. Hemez, Overview of uncertainty assessment for structural health monitoring, Tech. rep., Los Alamos National Lab., NM (US) (2001).

[79] Z. Lai, Y. Lei, S. Zhu, Y.-L. Xu, X.-H. Zhang, S. Krishnaswamy, Moving-window extended kalman filter for structural damage detection with unknown process and measurement noises, Measurement 88 (2016) 428 – 440.

[80] H. S. Park, Y. Kim, B. K. Oh, A model updating method with strain measurement from impact test for the safety of steel frame structures, Measurement 102 (2017) 220 – 229.

[81] T. Contursi, A. Messina, E. Williams, A multiple-damage location assurance criterion based on natural frequency changes, Journal of vibration and control 4 (5) (1998) 619–633.

[82] A. Messina, E. Williams, T. Contursi, Structural damage detection by a sensitivity and statistical-based method, Journal of sound and vibration 216 (5) (1998) 791–808.

[83] G. J. Yun, K. A. Ogorzalek, S. J. Dyke, W. Song, A parameter subset selection method using residual force vector for detecting multiple damage locations, Structural Control and Health Monitoring 17 (1) (2010) 48–67.

[84] S. Cao, H. Ouyang, Robust multi-damage localisation using common eigenvector analysis and covariance matrix changes, Mechanical Systems and Signal Processing 111 (2018) 663–677.

[85] H. Sohn, C. R. Farrar, N. F. Hunter, K. Worden, Structural health monitoring using statistical pattern recognition techniques, Journal of dynamic systems, measurement, and control 123 (4) (2001) 706–711.

[86] Z. Chen, C. Pan, L. Yu, Structural damage detection via adaptive dictionary learning and sparse representation of measured acceleration responses, Measurement 128 (2018) 377 – 387.

[87] C.-M. Chang, T.-K. Lin, C.-W. Chang, Applications of neural network models for structural health monitoring based on derived modal properties, Measurement 129

(2018) 457 – 470.

[88] M. Schwabacher, A survey of data-driven prognostics, in: Infotech@ Aerospace, 2005, p. 7002.

[89] Z.-H. Zhou, M.-L. Zhang, S.-J. Huang, Y.-F. Li, Multi-instance multi-label learning, Artificial Intelligence 176 (1) (2012) 2291–2320.

[90] P. Cawley, R. Adams, The location of defects in structures from measurements of natural frequencies, The Journal of Strain Analysis for Engineering Design 14 (2) (1979) 49–57.

[91] H. Lam, J. Ko, C. Wong, Localization of damaged structural connections based on experimental modal and sensitivity analysis, Journal of Sound and Vibration 210 (1) (1998) 91–115.

[92] Y. Yang, S. Nagarajaiah, Time-frequency blind source separation using independent component analysis for output-only modal identification of highly damped structures, Journal of Structural Engineering 139 (10) (2012) 1780–1793.

[93] S. W. Doebling, C. R. Farrar, M. B. Prime, D. W. Shevitz, Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review, Tech. rep., Los Alamos National Lab., NM (United States) (1996).

[94] R. Adams, P. Cawley, C. Pye, B. Stone, A vibration technique for non-destructively assessing the integrity of structures, Journal of Mechanical Engineering Science 20 (2) (1978) 93–100.

[95] Y.-S. Lee, M.-J. Chung, A study on crack detection using eigenfrequency test data, Computers & structures 77 (3) (2000) 327–342.

[96] J.-T. Kim, Y.-S. Ryu, H.-M. Cho, N. Stubbs, Damage identification in beam-type structures: frequency-based method vs mode-shape-based method, Engineering structures 25 (1) (2003) 57–67.

[97] J. R. Casas, A. C. Aparicio, Structural damage identification from dynamic-test data, Journal of Structural Engineering 120 (8) (1994) 2437–2450.

[98] A. Rytter, Vibrational based inspection of civil engineering structures, Ph.D. thesis, Aalborg University (1993).

[99] C. Farrar, G. James III, System identification from ambient vibration measurements on a bridge, Journal of Sound and Vibration 205 (1) (1997) 1–18.

[100] R. J. Allemang, D. L. Brown, A correlation coefficient for modal vector analysis, in: Proceedings of the 1st international modal analysis conference, Vol. 1, Orlando: Union College Press, 1982, pp. 110–116.

[101] J. Zhao, L. Zhang, Structural damage identification based on the modal data change, International Journal of Engineering and Manufacturing (IJEM) 2 (4) (2012) 59.

[102] J.-H. Kim, H.-S. Jeon, C.-W. Lee, Applications of the modal assurance criteria for detecting and locating structural faults, in: Proceedings of the International Modal Analysis Conference, SEM SOCIETY FOR EXPERIMENTAL MECHANICS INC, 1992, pp. 536–536.

[103] A. Pandey, M. Biswas, M. Samman, Damage detection from changes in curvature mode shapes, Journal of sound and vibration 145 (2) (1991) 321–332.

[104] M. A. Wahab, G. De Roeck, Damage detection in bridges using modal curvatures: application to a real damage scenario, Journal of Sound and vibration 226 (2) (1999) 217–235.

[105] D. M. Powers, Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation, Bioinfo Publications (2011).

[106] R. E. Schapire, Y. Singer, Boostexter: A boosting-based system for text categorization, Machine learning 39 (2-3) (2000) 135–168.

[107] T. Joachims, Text categorization with support vector machines: Learning with many relevant features, Machine learning: ECML-98 (1998) 137–142.

[108] Y. Yang, An evaluation of statistical approaches to text categorization, Information retrieval 1 (1) (1999) 69–90.

[109] M.-L. Zhang, Z.-H. Zhou, Multi-label learning by instance differentiation, in: AAAI, Vol. 7, 2007, pp. 669–674.

[110] J. J. Chen, C. Tsai, H. Moon, H. Ahn, J. J. Young, C.-h. Chen, The use of decision threshold adjustment in classification for cancer prediction (2006).

[111] F. Provost, T. Fawcett, R. Kohavi, The case against accuracy estimation for comparing classifiers. 5th int, in: Conference on Machine Learning, San Francisco, Kaufman Morgan, 1998.

[112] A. P. Bradley, The use of the area under the roc curve in the evaluation of machine learning algorithms, Pattern recognition 30 (7) (1997) 1145–1159.

[113] F. Provost, T. Fawcett, Robust classification for imprecise environments, Machine

learning 42 (3) (2001) 203–231.

[114] Y. Li, M. Zhang, W. Yang, Numerical and experimental investigation of modal-energy-based damage localization for offshore wind turbine structures, Advances in Structural Engineering 21 (10) (2018) 1510–1525.

[115] A. Z. Hosseinzadeh, G. G. Amiri, S. S. Razzaghi, K. Koo, S.-H. Sung, Structural damage detection using sparse sensors installation by optimization procedure based on the modal flexibility matrix, Journal of Sound and Vibration 381 (2016) 65–82.

[116] H. Li, D. Tao, Y. Huang, Y. Bao, A data-driven approach for seismic damage detection of shear-type building structures using the fractal dimension of time–frequency features, Structural Control and Health Monitoring 20 (9) (2013) 1191–1210.

[117] Z. Zhang, C. Sun, Multi-site structural damage identification using a multi-label classification scheme of machine learning, Measurement (2020) 107473.

[118] F. Poncelet, G. Kerschen, J.-C. Golinval, D. Verhelst, Output-only modal analysis using blind source separation techniques, Mechanical systems and signal processing 21 (6) (2007) 2335–2358.

[119] W. Zhou, D. Chelidze, Blind source separation based vibration mode identification, Mechanical systems and signal processing 21 (8) (2007) 3072–3087.

[120] G. Kerschen, F. Poncelet, J.-C. Golinval, Physical interpretation of independent component analysis in structural dynamics, Mechanical Systems and Signal Processing 21 (4) (2007) 1561–1575.

[121] A. Sadhu, S. Narasimhan, A decentralized blind source separation algorithm for ambient modal identification in the presence of narrowband disturbances, Structural Control and Health Monitoring 21 (3) (2014) 282–302.

[122] H. Sun, O. Büyüköztürk, The mit green building benchmark problem for structural health monitoring of tall buildings, Structural Control and Health Monitoring 25 (3) (2018) e2115.

[123] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, Neural networks 13 (4-5) (2000) 411–430.

[124] Y. Yang, S. Nagarajaiah, Blind identification of damage in time-varying systems using independent component analysis with wavelet transform, mechanical systems and signal processing 47 (1-2) (2014) 3–20.

[125] J. Antoni, Blind separation of vibration components: Principles and demonstrations, Mechanical Systems and Signal Processing 19 (6) (2005) 1166–1180.

[126] Q. He, Z. Feng, F. Kong, Detection of signal transients using independent component analysis and its application in gearbox condition monitoring, Mechanical systems and signal processing 21 (5) (2007) 2056–2071.

[127] H. Song, L. Zhong, B. Han, Structural damage detection by integrating independent component analysis and support vector machine, International journal of systems science 37 (13) (2006) 961–967.

[128] C. Zang, M. I. Friswell, M. Imregun, Structural damage detection using independent component analysis, Structural Health Monitoring 3 (1) (2004) 69–83.

[129] M. Salehi, S. Ziaei-Rad, M. Ghayour, M. A. Vaziri-Zanjani, A frequency response based structural damage localization method using independent component analysis, Journal of Mechanical Science and Technology 27 (3) (2013) 609–619.

[130] J. Wodecki, P. Stefaniak, M. Sawicki, R. Zimroz, Application of independent component analysis in temperature data analysis for gearbox fault detection, in: Cyclostationarity: Theory and Methods III, Springer, 2017, pp. 187–198.

[131] A. Mahvash, A. A. Lakis, Independent component analysis as applied to vibration source separation and fault diagnosis, Journal of Vibration and Control 22 (6) (2016) 1682–1692.

[132] C. J. James, O. J. Gibson, Temporally constrained ica: an application to artifact rejection in electromagnetic brain signal analysis, IEEE Transactions on Biomedical Engineering 50 (9) (2003) 1108–1116.

[133] W. Lu, J. C. Rajapakse, Ica with reference, Neurocomputing 69 (16-18) (2006) 2244–2257.

[134] G. Yu, Fault feature extraction using independent component analysis with reference and its application on fault diagnosis of rotating machinery, Neural computing and applications 26 (1) (2015) 187–198.

[135] Z. Wang, J. Chen, G. Dong, Y. Zhou, Constrained independent component analysis and its application to machine fault diagnosis, Mechanical Systems and Signal Processing 25 (7) (2011) 2501–2512.

[136] A. Hyvärinen, J. Karhunen, E. Oja, Independent component analysis, Vol. 46, John Wiley & Sons, 2004.

[137] M. B. Christopher, PATTERN RECOGNITION AND MACHINE LEARNING., Springer-Verlag New York, 2016.

[138] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometrics and

intelligent laboratory systems 2 (1-3) (1987) 37–52.

[139] K. N. Kesavan, A. S. Kiremidjian, A wavelet-based damage diagnosis algorithm using principal component analysis, Structural Control and Health Monitoring 19 (8) (2012) 672–685.

[140] D. Tibaduiza, L. Mujica, J. Rodellar, Damage classification in structural health monitoring using principal component analysis and self-organizing maps, Structural Control and Health Monitoring 20 (10) (2013) 1303–1316.

[141] R. De Maesschalck, D. Jouan-Rimbaud, D. L. Massart, The mahalanobis distance, Chemometrics and intelligent laboratory systems 50 (1) (2000) 1–18.

[142] K. Q. Weinberger, L. K. Saul, Distance metric learning for large margin nearest neighbor classification, Journal of Machine Learning Research 10 (Feb) (2009) 207–244.

[143] S. Xiang, F. Nie, C. Zhang, Learning a mahalanobis distance metric for data clustering and classification, Pattern Recognition 41 (12) (2008) 3600–3612.

[144] D. Wang, D. S. Yeung, E. C. Tsang, Weighted mahalanobis distance kernels for support vector machines, IEEE Transactions on Neural Networks 18 (5) (2007) 1453–1462.

[145] S. Theodoridis, K. Koutroumbas, Chapter 5 - feature selection, in: S. Theodoridis, K. Koutroumbas (Eds.), Pattern Recognition (Fourth Edition), fourth edition Edition, Academic Press, Boston, 2009, pp. 261 – 322. doi:https://doi.org/10.1016/B978-1-59749-272-0.50007-4.

[146] K. Worden, G. Manson, N. R. Fieller, Damage detection using outlier analysis, Journal of Sound and Vibration 229 (3) (2000) 647–667.

[147] Global wind report, Global Wind Energy Council, 2019.

[148] S. Soua, P. Van Lieshout, A. Perera, T.-H. Gan, B. Bridge, Determination of the combined vibrational and acoustic emission signature of a wind turbine gearbox and generator shaft in service as a pre-requisite for effective condition monitoring, Renewable Energy 51 (2013) 175–181.

[149] A. Kolios, A. Chahardehi, F. Brennan, Experimental determination of the overturning moment and net lateral force generated by a novel vertical axis wind turbine: experiment design under load uncertainty, Experimental Techniques 37 (1) (2013) 7–14.

[150] F. Brennan, A. Kolios, Structural integrity considerations for the h2ocean multi

modal wind-wave platform, in: European Wind Energy Association Conference and Exhibition 2014, EWEA 2014, European Wind Energy Association, 2014.

[151] M. Borg, M. Collu, A. Kolios, Offshore floating vertical axis wind turbines, dynamics modelling state of the art. part ii: Mooring line and structural dynamics, Renewable and Sustainable Energy Reviews 39 (2014) 1226–1234.

[152] J. D. Sørensen, Framework for risk-based planning of operation and maintenance for offshore wind turbines, Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology 12 (5) (2009) 493–506.

[153] M. Martinez-Luengo, A. Kolios, L. Wang, Structural health monitoring of offshore wind turbines: A review through the statistical pattern recognition paradigm, Renewable and Sustainable Energy Reviews 64 (2016) 91–105.

[154] D. E. Bently, T. Hatch'Charles, Fundamentals of rotating machinery diagnostics, Mechanical Engineering-CIME 125 (12) (2003) 53–54.

[155] P. J. Shull, Nondestructive evaluation: theory, techniques, and applications, CRC press, 2002.

[156] C. C. Ciang, J.-R. Lee, H.-J. Bang, Structural health monitoring for a wind turbine system: a review of damage detection methods, Measurement science and technology 19 (12) (2008) 122001.

[157] W. Liu, B. Tang, J. Han, X. Lu, N. Hu, Z. He, The structure healthy condition monitoring and fault diagnosis methods in wind turbines: A review, Renewable and Sustainable Energy Reviews 44 (2015) 466–472.

[158] D. Li, S.-C. M. Ho, G. Song, L. Ren, H. Li, A review of damage detection methods for wind turbine blades, Smart Materials and Structures 24 (3) (2015) 033001.

[159] M. Martinez-Luengo, M. Shafiee, Guidelines and cost-benefit analysis of the structural health monitoring implementation in offshore wind turbine support structures, Energies 12 (6) (2019) 1176.

[160] M. Martinez-Luengo, M. Shafiee, A. Kolios, Data management for structural integrity assessment of offshore wind turbine support structures: data cleansing and missing data imputation, Ocean Engineering 173 (2019) 867–883.

[161] M. Ozbek, F. Meng, D. J. Rixen, Challenges in testing and monitoring the in-operation vibration characteristics of wind turbines, Mechanical Systems and Signal Processing 41 (1-2) (2013) 649–666.

[162] W.-H. Hu, S. Thöns, R. G. Rohrmann, S. Said, W. Rücker, Vibration-based struc-

tural health monitoring of a wind turbine system. part i: Resonance phenomenon, Engineering Structures 89 (2015) 260–272.

[163] W.-H. Hu, S. Thöns, R. G. Rohrmann, S. Said, W. Rücker, Vibration-based structural health monitoring of a wind turbine system part ii: Environmental/operational effects on dynamic properties, Engineering Structures 89 (2015) 273–290.

[164] C. Devriendt, F. Magalhães, W. Weijtjens, G. De Sitter, Á. Cunha, P. Guillaume, Structural health monitoring of offshore wind turbines using automated operational modal analysis, Structural Health Monitoring 13 (6) (2014) 644–659.

[165] A. Bajrić, J. Høgsberg, F. Rüdinger, Evaluation of damping estimates by automated operational modal analysis for offshore wind turbine tower vibrations, Renewable Energy 116 (2018) 153–163.

[166] P. Joosse, M. Blanch, A. Dutton, D. Kouroussis, T. Philippidis, P. Vionis, Acoustic emission monitoring of small wind turbine blades, J. Sol. Energy Eng. 124 (4) (2002) 446–454.

[167] R. Raišutis, E. Jasiūnienė, E. Žukauskas, Ultrasonic ndt of wind turbine blades using guided waves, Ultragarsas 63 (1) (2008) 7–11.

[168] L. J. Prendergast, K. Gavin, P. Doherty, An investigation into the effect of scour on the natural frequency of an offshore wind turbine, Ocean Engineering 101 (2015) 1–11.

[169] H.-J. Bang, H.-I. Kim, K.-S. Lee, Measurement of strain and bending deflection of a wind turbine tower using arrayed fbg sensors, International journal of precision engineering and manufacturing 13 (12) (2012) 2121–2126.

[170] A. Baldassarre, A. Ceruti, D. N. Valyou, P. Marzocca, Towards a digital twin realization of the blade system design study wind turbine blade, Wind and Structures 28 (5) (2019) 271–284.

[171] E. A. Camargo, J. P. Ulfkjaer, R. Brincker, J. Nøergaard, S. S. Gadegaard, Operational modal analysis and finite-element model updating of pilot concrete wind turbine tower, Journal of Structural Engineering 145 (2) (2018) 05018003.

[172] J. Lin, L. K. Leung, Y.-L. Xu, S. Zhan, S. Zhu, Field measurement, model updating, and response prediction of a large-scale straight-bladed vertical axis wind turbine structure, Measurement 130 (2018) 57–70.

[173] M. Luczak, S. Manzato, B. Peeters, K. Branner, P. Berring, M. Kahsin, Updating finite element model of a wind turbine blade section using experimental modal analysis results, Shock and Vibration 2014.

[174] Y. Xu, G. Nikitas, T. Zhang, Q. Han, M. Chryssanthopoulos, S. Bhattacharya, Y. Wang, Support condition monitoring of offshore wind turbines using model updating techniques, Structural Health Monitoring (2019) 1475921719875628.

[175] C. Sun, Mitigation of offshore wind turbine responses under wind and wave loading: Considering soil effects and damage, Structural Control and Health Monitoring 25 (3) (2018) e2117.

[176] C. Sun, Semi-active control of monopile offshore wind turbines under multi-hazards, Mechanical Systems and Signal Processing 99 (2018) 285–305.

[177] C. Sun, V. Jahangiri, Bi-directional vibration control of offshore wind turbines using a 3d pendulum tuned mass damper, Mechanical Systems and Signal Processing 105 (2018) 338–360.

[178] C. Sun, V. Jahangiri, Fatigue damage mitigation of offshore wind turbines under real wind and wave conditions, Engineering Structures 178 (2019) 472–483.

[179] V. Jahangiri, C. Sun, Integrated bi-directional vibration control and energy harvesting of monopile offshore wind turbines, Ocean Engineering 178 (2019) 260–269.

[180] D. Zhu, X. Dong, Y. Wang, Substructure stiffness and mass updating through minimization of modal dynamic residuals, Journal of Engineering Mechanics 142 (5) (2016) 04016013.

[181] R. Brincker, L. Zhang, P. Andersen, Modal identification from ambient responses using frequency domain decomposition, in: Proc. of the 18*'International Modal Analysis Conference (IMAC), San Antonio, Texas, 2000.

Vita

Zhiming Zhang was born in 1988 in Hebei Province, People's Republic of China. He finished his undergraduate studies at Liaoning Technical University, People's Republic of China, in July 2010. He earned a master of science degree in bridge engineering from Southwest Jiaotong University in July 2013. In August 2016 he joined Louisiana State University to pursue graduate studies in civil engineering with a focus on structural engineering. He is currently a candidate for the degree of Doctor of Philosophy in civil engineering, which is anticipated to be awarded in August 2020.